AM

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47	A2	(11) International Publication Number: WO 99/06548 (43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/IB98/01222

(22) International Filing Date: 31 July 1998 (31.07.98)

(30) Priority Data:
08/905,135 1 August 1997 (01.08.97) US

(71) Applicant (*for all designated States except US*): GENSET [FR/FR]; 24, rue Royale, F-75008 Paris (FR).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DUMAS MILNE EDWARDS, Jean-Baptiste [FR/FR]; 8, rue Grégoire-de-Tours, F-75006 Paris (FR). DUCLERT, Aymeric [FR/FR]; 6 ter, rue Victorine, F-94100 Saint-Maur (FR). LACROIX, Bruno [FR/FR]; 93, route de Vourles, F-69230 Saint-Genis Laval (FR).

(74) Agents: MARTIN, Jean-Jacques et al.; Cabinet Régimbeau, 26, Avenue Kléber, F-75116 Paris (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: 5'ESTs FOR NON TISSUE SPECIFIC SECRETED PROTEINS

(57) Abstract

The sequences of 5'ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5'ESTs may be to obtain cDNAs and genomic DNAs corresponding to the 5'ESTs. The 5'ESTs may also be used in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5'ESTs. The 5'ESTs may also be used to design expression vectors and secretion vectors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

5' ESTs FOR NON TISSUE SPECIFIC SECRETED PROTEINS

Background of the Invention

The estimated 50,000-100,000 genes scattered along the human chromosomes offer tremendous promise for the understanding, diagnosis, and treatment of human diseases. In addition, probes capable of specifically hybridizing to loci distributed throughout the human genome find applications in the construction of high resolution chromosome maps and in the identification of individuals.

In the past, the characterization of even a single human gene was a painstaking process, requiring years of effort. Recent developments in the areas of cloning vectors, DNA sequencing, and computer technology have merged to greatly accelerate the rate at which human genes can be isolated, sequenced, mapped, and characterized. Cloning vectors such as yeast artificial chromosomes (YACs) and bacterial artificial chromosomes (BACs) are able to accept DNA inserts ranging from 300 to 1000 kilobases (kb) or 100-400 kb in length respectively, thereby facilitating the manipulation and ordering of DNA sequences distributed over great distances on the human chromosomes. Automated DNA sequencing machines permit the rapid sequencing of human genes. Bioinformatics software enables the comparison of nucleic acid and protein sequences, thereby assisting in the characterization of human gene products.

Currently, two different approaches are being pursued for identifying and characterizing the genes distributed along the human genome. In one approach, large fragments of genomic DNA are isolated, cloned, and sequenced. Potential open reading frames in these genomic sequences are identified using bioinformatics software. However, this approach entails sequencing large stretches of human DNA which do not encode proteins in order to find the protein encoding sequences scattered throughout the genome. In addition to requiring extensive sequencing, the bioinformatics software may mischaracterize the genomic sequences obtained. Thus, the software may produce false positives in which non-coding DNA is mischaracterized as coding DNA or false negatives in which coding DNA is mislabeled as non-coding DNA.

An alternative approach takes a more direct route to identifying and characterizing human genes. In this approach, complementary DNAs (cDNAs) are synthesized from isolated messenger RNAs (mRNAs) which encode human proteins. Using this approach,

sequencing is only performed on DNA which is derived from protein coding portions of the genome. Often, only short stretches of the cDNAs are sequenced to obtain sequences called expressed sequence tags (ESTs). The ESTs may then be used to isolate or purify extended cDNAs which include sequences adjacent to the EST sequences. The extended cDNAs may contain all of the sequence of the EST which was used to obtain them or only a portion of the sequence of the EST which was used to obtain them. In addition, the extended cDNAs may contain the full coding sequence of the gene from which the EST was derived or, alternatively, the extended cDNAs may include portions of the coding sequence of the gene from which the EST was derived. It will be appreciated that there may be several extended cDNAs which include the EST sequence as a result of alternate splicing or the activity of alternative promoters.

In the past, these short EST sequences were often obtained from oligo-dT primed cDNA libraries. Accordingly, they mainly corresponded to the 3′ untranslated region of the mRNA. In part, the prevalence of EST sequences derived from the 3′ end of the mRNA is a result of the fact that typical techniques for obtaining cDNAs are not well suited for isolating cDNA sequences derived from the 5′ ends of mRNAs. (Adams *et al., Nature* 377:3-174, 1996; Hillier *et al., Genome Res.* 6:807-828, 1996).

In addition, in those reported instances where longer cDNA sequences have been obtained, the reported sequences typically correspond to coding sequences and do not include the full 5′ untranslated region of the mRNA from which the cDNA is derived. Such incomplete sequences may not include the first exon of the mRNA, particularly in situations where the first exon is short. Furthermore, they may not include some exons, often short ones, which are located upstream of splicing sites. Thus, there is a need to obtain sequences derived from the 5′ ends of mRNAs.

While many sequences derived from human chromosomes have practical applications, approaches based on the identification and characterization of those chromosomal sequences which encode a protein product are particularly relevant to diagnostic and therapeutic uses. Of the 50,000-100,000 protein coding genes, those genes encoding proteins which are secreted from the cell in which they are synthesized, as well as the secreted proteins themselves, are particularly valuable as potential therapeutic agents. Such proteins are often

involved in cell to cell communication and may be responsible for producing a clinically relevant response in their target cells.

In fact, several secretory proteins, including tissue plasminogen activator, G-CSF, GM-CSF, erythropoietin, human growth hormone, insulin, interferon-α, interferon-β, interferon-γ, and interleukin-2, are currently in clinical use. These proteins are used to treat a wide range of conditions, including acute myocardial infarction, acute ischemic stroke, anemia, diabetes, growth hormone deficiency, hepatitis, kidney carcinoma, chemotherapy induced neutropenia and multiple sclerosis. For these reasons, extended cDNAs encoding secreted proteins or portions thereof represent a particularly valuable source of therapeutic agents. Thus, there is a need for the identification and characterization of secreted proteins and the nucleic acids encoding them.

In addition to being therapeutically useful themselves, secretory proteins include short peptides, called signal peptides, at their amino termini which direct their secretion. These signal peptides are encoded by the signal sequences located at the 5′ ends of the coding sequences of genes encoding secreted proteins. Because these signal peptides will direct the extracellular secretion of any protein to which they are operably linked, the signal sequences may be exploited to direct the efficient secretion of any protein by operably linking the signal sequences to a gene encoding the protein for which secretion is desired. In addition, portions of signal sequences may also be used to direct the intracellular import of a peptide or protein of interest. This may prove beneficial in gene therapy strategies in which it is desired to deliver a particular gene product to cells other than the cell in which it is produced. Signal sequences encoding signal peptides also find application in simplifying protein purification techniques. In such applications, the extracellular secretion of the desired protein greatly facilitates purification by reducing the number of undesired proteins from which the desired protein must be selected. Thus, there exists a need to identify and characterize the 5′ portions of the genes for secretory proteins which encode signal peptides.

Public information on the number of human genes for which the promoters and upstream regulatory regions have been identified and characterized is quite limited. In part, this may be due to the difficulty of isolating such regulatory sequences. Upstream regulatory sequences such as transcription factor binding sites are typically too short to be utilized as probes for isolating promoters from human genomic libraries. Recently, some approaches

have been developed to isolate human promoters. One of them consists of making a CpG island library (Cross, *et al., Nature Genetics* 6: 236-244, 1994). The second consists of isolating human genomic DNA sequences containing SpeI binding sites by the use of SpeI binding protein. (Mortlock *et al., Genome Res.* 6:327-335, 1996). Both of these approaches have their limits due to a lack of specificity or of comprehensiveness.

The present 5′ ESTs may be used to efficiently identify and isolate upstream regulatory regions which control the location, developmental stage, rate, and quantity of protein synthesis, as well as the stability of the mRNA. (Theil, *BioFactors* 4:87-93, 1993). Once identified and characterized, these regulatory regions may be utilized in gene therapy or protein purification schemes to obtain the desired amount and locations of protein synthesis or to inhibit, reduce, or prevent the synthesis of undesirable gene products.

In addition, ESTs containing the 5′ ends of secretory protein genes may include sequences useful as probes for chromosome mapping and the identification of individuals. Thus, there is a need to identify and characterize the sequences upstream of the 5′ coding sequences of genes encoding secretory proteins.

Summary of the Invention

The present invention relates to purified, isolated, or recombinant ESTs which include sequences derived from the authentic 5′ ends of their corresponding mRNAs. The term "corresponding mRNA" refers to the mRNA which was the template for the cDNA synthesis which produced the 5′ EST. These sequences will be referred to hereinafter as "5′ ESTs." As used herein, the term "purified" does not require absolute purity; rather, it is intended as a relative definition. Individual 5′ EST clones isolated from a cDNA library have been conventionally purified to electrophoretic homogeneity. The sequences obtained from these clones could not be obtained directly either from the library or from total human DNA. The cDNA clones are not naturally occurring as such, but rather are obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The conversion of mRNA into a cDNA library involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection. Thus, creating a cDNA library from messenger RNA and subsequently isolating individual clones from that library results in an approximately 10^4-10^6 fold purification of the native message.

Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated.

As used herein, the term "isolated" requires that the material be removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide present in a living animal is not isolated, but the same polynucleotide, separated from some or all of the coexisting materials in the natural system, is isolated.

As used herein, the term "recombinant" means that the 5′ EST is adjacent to "backbone" nucleic acid to which it is not adjacent in its natural environment. Additionally, to be "enriched" the 5′ ESTs will represent 5% or more of the number of nucleic acid inserts in a population of nucleic acid backbone molecules. Backbone molecules according to the present invention include nucleic acids such as expression vectors, self-replicating nucleic acids, viruses, integrating nucleic acids, and other vectors or nucleic acids used to maintain or manipulate a nucleic acid insert of interest. Preferably, the enriched 5′ ESTs represent 15% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. More preferably, the enriched 5′ ESTs represent 50% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules. In a highly preferred embodiment, the enriched 5′ ESTs represent 90% or more of the number of nucleic acid inserts in the population of recombinant backbone molecules.

"Stringent", moderate," and "low" hybridization conditions are as defined in Example 29.

Unless otherwise indicated, a "complementary" sequence is fully complementary.

Thus, 5′ ESTs in cDNA libraries in which one or more 5′ ESTs make up 5% or more of the number of nucleic acid inserts in the backbone molecules are "enriched recombinant 5′ ESTs" as defined herein. Likewise, 5′ ESTs in a population of plasmids in which one or more 5′ EST of the present invention have been inserted such that they represent 5% or more of the number of inserts in the plasmid backbone are " enriched recombinant 5′ ESTs" as defined herein. However, 5′ ESTs in cDNA libraries in which 5′ ESTs constitute less than 5% of the number of nucleic acid inserts in the population of backbone molecules, such as libraries in

which backbone molecules having a 5′ EST insert are extremely rare, are not "enriched recombinant 5′ ESTs."

In particular, the present invention relates to 5′ ESTs which are derived from genes encoding secreted proteins. As used herein, a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal peptides in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g. soluble proteins), or partially (e.g. receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Such 5′ ESTs include nucleic acid sequences, called signal sequences, which encode signal peptides which direct the extracellular secretion of the proteins encoded by the genes from which the 5′ ESTs are derived. Generally, the signal peptides are located at the amino termini of secreted proteins.

Secreted proteins are translated by ribosomes associated with the "rough" endoplasmic reticulum. Generally, secreted proteins are co-translationally transferred to the membrane of the endoplasmic reticulum. Association of the ribosome with the endoplasmic reticulum during translation of secreted proteins is mediated by the signal peptide. The signal peptide is typically cleaved following its co-translational entry into the endoplasmic reticulum. After delivery to the endoplasmic reticulum, secreted proteins may proceed through the Golgi apparatus. In the Golgi apparatus, the proteins may undergo post-translational modification before entering secretory vesicles which transport them across the cell membrane.

The 5′ ESTs of the present invention have several important applications. For example, they may be used to obtain and express cDNA clones which include the full protein coding sequences of the corresponding gene products, including the authentic translation start sites derived from the 5′ ends of the coding sequences of the mRNAs from which the 5′ ESTs are derived. These cDNAs will be referred to hereinafter as "full length cDNAs." These cDNAs may also include DNA derived from mRNA sequences upstream of the translation start site. The full length cDNA sequences may be used to express the proteins corresponding to the 5′ ESTs. As discussed above, secreted proteins are therapeutically important. Thus, the proteins expressed from the cDNAs may be useful in treating or

controlling a variety of human conditions. The 5′ ESTs may also be used to obtain the corresponding genomic DNA. The term "corresponding genomic DNA" refers to the genomic DNA which encodes the mRNA from which the 5′ EST was derived.

Alternatively, the 5′ ESTs may be used to obtain and express extended cDNAs encoding portions of the secreted protein. The portions may comprise the signal peptides of the secreted proteins or the mature proteins generated when the signal peptide is cleaved off. The portions may also comprise polypeptides having at least 10 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. Alternatively, the portions may comprise at least 15 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. In some embodiments, the portions may comprise at least 25 consecutive amino acids encoded by the extended cDNAs or full length cDNAs. In other embodiments, the portions may comprise at least 40 amino acids encoded by the extended cDNAs or full length cDNAs.

Antibodies which specifically recognize the entire secreted proteins encoded by the extended cDNAs, full length cDNAs, or fragments thereof having at least 10 consecutive amino acids, at least 15 consecutive amino acids, at least 25 consecutive amino acids, or at least 40 consecutive amino acids may also be obtained as described below. Antibodies which specifically recognize the mature protein generated when the signal peptide is cleaved may also be obtained as described below. Similarly, antibodies which specifically recognize the signal peptides encoded by the extended cDNAs or full length cDNAs may also be obtained.

In some embodiments, the extended cDNAs obtained using the 5′ ESTs include the signal sequence. In other embodiments, the extended cDNAs obtained using the 5′ ESTs may include the full coding sequence for the mature protein (*i.e.* the protein generated when the signal polypeptide is cleaved off). In addition, the extended cDNAs obtained using the 5′ ESTs may include regulatory regions upstream of the translation start site or downstream of the stop codon which control the amount, location, or developmental stage of gene expression.

As discussed above, secreted proteins are therapeutically important. Thus, the proteins expressed from the extended cDNAs or full length cDNAs obtained using the 5′ ESTs may be useful in treating or controlling a variety of human conditions.

The 5′ ESTs (or cDNAs or genomic DNAs obtained therefrom) may be used in forensic procedures to identify individuals or in diagnostic procedures to identify individuals having genetic diseases resulting from abnormal expression of the genes corresponding to the 5′ ESTs. In addition, the present invention is useful for constructing a high resolution map of the human chromosomes.

The present invention also relates to secretion vectors capable of directing the secretion of a protein of interest. Such vectors may be used in gene therapy strategies in which it is desired to produce a gene product in one cell which is to be delivered to another location in the body. Secretion vectors may also facilitate the purification of desired proteins.

The present invention also relates to expression vectors capable of directing the expression of an inserted gene in a desired spatial or temporal manner or at a desired level. Such vectors may include sequences upstream of the 5′ ESTs, such as promoters or upstream regulatory sequences.

Finally, the present invention may also be used for gene therapy to control or treat genetic diseases. Signal peptides may also be fused to heterologous proteins to direct their extracellular secretion.

Bacterial clones containing Bluescript plasmids having inserts containing the 5′ ESTs of the present invention (SEQ ID NOs: 38-291 are presently stored at 80°C in 4% (v/v) glycerol in the inventor's laboratories under the designations listed next to the SEQ ID NOs in II). The inserts may be recovered from the deposited materials by growing the appropriate clones on a suitable medium. The Bluescript DNA can then be isolated using plasmid isolation procedures familiar to those skilled in the art such as alkaline lysis minipreps or large scale alkaline lysis plasmid isolation procedures. If desired the plasmid DNA may be further enriched by centrifugation on a cesium chloride gradient, size exclusion chromatography, or anion exchange chromatography. The plasmid DNA obtained using these procedures may then be manipulated using standard cloning techniques familiar to those skilled in the art. Alternatively, a PCR can be done with primers designed at both ends of the EST insertion. The PCR product which corresponds to the 5′ EST can then be manipulated using standard cloning techniques familiar to those skilled in the art.

One aspect of the present invention is a purified or isolated nucleic acid having the sequence of one of SEQ ID NOs: 38-291 or having a sequence complementary thereto. In one embodiment, the nucleic acid is recombinant.

Another aspect of the present invention is a purified or isolated nucleic acid comprising at least 10 consecutive bases of the sequence of one of SEQ ID NOs: 38-291 or one of the sequences complementary thereto.

Yet another aspect of the present invention is a purified or isolated nucleic acid comprising at least 15 consecutive bases of one of the sequences of SEQ ID NOs: 38-291 or one of the sequences complementary thereto. In one embodiment, the nucleic acid is recombinant.

A further aspect of the present invention is a purified or isolated nucleic acid of at least 15 bases capable of hybridizing under stringent conditions to the sequence of one of SEQ ID NOs: 38-291 or one of the sequences complementary to the sequences of SEQ ID NOs: 38-291. In one embodiment, the nucleic acid is recombinant.

Another aspect of the present invention is a purified or isolated nucleic acid encoding a human gene product, said human gene product having a sequence partially encoded by one of the sequences of SEQ ID NO: 38-291.

Still another aspect of the present invention is a method of making a cDNA encoding a human secretory protein, said human secretory protein being partially encoded by one of SEQ ID NOs 38-291, comprising the steps of contacting a collection of mRNA molecules from human cells with a primer comprising at least 15 consecutive nucleotides of a sequence complementary to one of SEQ ID NOs: 38-291; hybridizing said primer to an mRNA in said collection that encodes said protein; reverse transcribing said hybridized primer to make a first cDNA strand from said mRNA; making a second cDNA strand complementary to said first cDNA strand; and isolating the resulting cDNA encoding said protein comprising said first cDNA strand and said second cDNA strand.

Another aspect of the invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the

cDNA comprises the full protein coding sequence of said protein which sequence is partially included in one of the sequences of SEQ ID NOs: 38-291.

Another aspect of the present invention is a method of making a cDNA encoding a human secretory protein that is partially encoded by one of SEQ ID NOs 38-291, comprising the steps of obtaining a cDNA comprising one of the sequences of SEQ ID NOs: 38-291; contacting said cDNA with a detectable probe comprising at least 15 consecutive nucleotides of said sequence of SEQ ID NO: 38-291 or a sequence complementary thereto under conditions which permit said probe to hybridize to said cDNA; identifying a cDNA which hybridizes to said detectable probe; and isolating said cDNA which hybridizes to said probe.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-291.

Another aspect of the present invention is a method of making a cDNA comprising one of the sequence of SEQ ID NOs: 38-291, comprising the steps of contacting a collection of mRNA molecules from human cells with a first primer capable of hybridizing to the polyA tail of said mRNA; hybridizing said first primer to said polyA tail; reverse transcribing said mRNA to make a first cDNA strand; making a second cDNA strand complementary to said first cDNA strand using at least one primer comprising at least 15 nucleotides of one of the sequences of SEQ ID NOs 38-291; and isolating the resulting cDNA comprising said first cDNA strand and said second cDNA strand.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-291.

In one embodiment of the method described in the two paragraphs above, the second cDNA strand is made by contacting said first cDNA strand with a first pair of primers, said

first pair of primers comprising a second primer comprising at least 15 consecutive nucleotides of one of the sequences of SEQ ID NOs 38-291 and a third primer having a sequence therein which is included within the sequence of said first primer; performing a first polymerase chain reaction with said first pair of nested primers to generate a first PCR product; contacting said first PCR product with a second pair of primers, said second pair of primers comprising a fourth primer, said fourth primer comprising at least 15 consecutive nucleotides of said sequence of one of SEQ ID NOs: 38-291 , and a fifth primer, said fourth and fifth primers being capable of hybridizing to sequences within said first PCR product; and performing a second polymerase chain reaction, thereby generating a second PCR product.

One aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291, or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-291.

Another aspect of the present invention is the method described four paragraphs above in which the second cDNA strand is made by contacting said first cDNA strand with a second primer comprising at least 15 consecutive nucleotides of the sequences of SEQ ID NOs: 38-291; hybridizing said second primer to said first strand cDNA; and extending said hybridized second primer to generate said second cDNA strand.

Another aspect of the present invention is an isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein partially encoded by one of SEQ ID NOs 38-291 or comprising a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method described in the preceding paragraph. In one embodiment, the cDNA comprises the full protein coding sequence partially included in of one of the sequences of SEQ ID NOs: 38-291.

Another aspect of the present invention is a method of making a protein comprising one of the sequences of SEQ ID NOs: 292-545, comprising the steps of obtaining a cDNA encoding the full protein sequence partially included in one of the sequences of sequence of SEQ ID NOs: 38-291; inserting said cDNA in an expression vector such that said cDNA is

operably linked to a promoter; introducing said expression vector into a host cell whereby said host cell produces the protein encoded by said cDNA; and isolating said protein.

Another aspect of the present invention is an isolated protein obtainable by the method described in the preceding paragraph.

Another aspect of the present invention is a method of obtaining a promoter DNA comprising the steps of obtaining DNAs located upstream of the nucleic acids of SEQ ID NOs: 38-291 or the sequences complementary thereto; screening said upstream DNAs to identify a promoter capable of directing transcription initiation; and isolating said DNA comprising said identified promoter. In one embodiment, the obtaining step comprises chromosome walking from said nucleic acids of SEQ ID NOs: 38-291 or sequences complementary thereto. In another embodiment, the screening step comprises inserting said upstream sequences into a promoter reporter vector. In another embodiment, the screening step comprises identifying motifs in said upstream DNAs which are transcription factor binding sites or transcription start sites.

Another aspect of the present invention is an isolated promoter obtainable by the method described above.

Another aspect of the present invention is an isolated or purified protein comprising one of the sequences of SEQ ID NOs: 292-545.

Another aspect of the present invention is the inclusion of at least one of the sequences of SEQ ID NOs: 38-291, or one of the sequences complementary to the sequences of SEQ ID NOs: 38-291, or a fragment thereof of at least 15 consecutive nucleotides in an array of discrete ESTs or fragments thereof of at least 15 nucleotides in length. In one embodiment, the array includes at least two of the sequences of SEQ ID NOs: 38-291, the sequences complementary to the sequences of SEQ ID NOs: 38-291, or fragments thereof of at least 15 consecutive nucleotides. In another embodiment, the array includes at least five of the sequences of SEQ ID NOs: 38-291, the sequences complementary to the sequences of SEQ ID NOs: 38-291, or fragments thereof of at least 15 consecutive nucleotides.

Another aspect of the present invention is a promoter having a sequence selected from the group consisting of SEQ ID NOs: 31, 34, and 37.

Brief Description of the Drawings

Figure 1 is a summary of a procedure for obtaining cDNAs which have been selected to include the 5′ ends of the mRNAs from which they derived.

Figure 2 shows the distribution of Von Heijne scores for 5′ ESTs in each of the categories described herein and the probability that these 5′ ESTs encode a signal peptide.

Figure 3 summarizes a general method used to clone and sequence extended cDNAs containing sequences adjacent to 5' ESTs.

Figure 4 (description of promoters structure isolated from SignalTag 5' ESTs) provides a schematic description of promoters isolated and the way they are assembled with the corresponding 5′ tags.

Detailed Description of the Preferred Embodiment

Table IV is an analysis of the 43 amino acids located at the N terminus of all human SwissProt proteins to determine the frequency of false positives and false negatives using the techniques for signal peptide identification described herein.

Table V shows the distribution of 5′ ESTs in each category described herein and the number of 5′ ESTs in each category having a given minimum Von Heijne's score.

Table VI shows the distribution of 5′ ESTs in each category described herein with respect to the tissue from which the 5' ESTs of the corresponding mRNA were obtained.

Table VII describes the transcription factor binding sites present in each of these promoters.

I. General Methods for Obtaining 5′ ESTs derived from mRNAs with intact 5' ends

In order to obtain the 5' ESTs of the present invention, mRNAs with intact 5' ends must be obtained. Currently, there are two approaches for obtaining such mRNAs with intact 5' ends as described below: either chemical (1) or enzymatic (2).

1. Chemical Methods for Obtaining mRNAs having Intact 5′ Ends

One of these approaches is a chemical modification method involving derivatization of the 5′ ends of the mRNAs and selection of the derivatized mRNAs. The 5′ ends of

eukaryotic mRNAs possess a structure referred to as a "cap" which comprises a guanosine methylated at the 7 position. The cap is joined to the first transcribed base of the mRNA by a 5′, 5′-triphosphate bond. In some instances, the 5′ guanosine is methylated in both the 2 and 7 positions. Rarely, the 5′ guanosine is trimethylated at the 2, 7 and 7 positions. In the chemical method for obtaining mRNAs having intact 5′ ends, the 5′ cap is specifically derivatized and coupled to a reactive group on an immobilizing substrate. This specific derivatization is based on the fact that only the ribose linked to the methylated guanosine at the 5′ end of the mRNA and the ribose linked to the base at the 3′ terminus of the mRNA, possess 2′, 3′-cis diols.

Optionally, the 2′, 3′-cis diol of the 3′ terminal ribose may be chemically modified, substituted, converted, or eliminated, leaving only the ribose linked to the methylated guanosine at the 5′ end of the mRNA with a 2′, 3′-cis diol. A variety of techniques are available for eliminating the 2′, 3′-cis diol on the 3′ terminal ribose. For example, controlled alkaline hydrolysis may be used to generate mRNA fragments in which the 3′ terminal ribose is a 3′-phosphate, 2′-phosphate or (2′, 3′)-cyclophosphate. Thereafter, the fragment which includes the original 3′ ribose may be eliminated from the mixture through chromatography on an oligodT column. Alternatively, a base which lacks the 2′, 3′-cis diol may be added to the 3′ end of the mRNA using an RNA ligase such as T4 RNA ligase. Example 1 below describes a method for ligation of a nucleoside diphosphate to the 3′ end of messenger RNA.

EXAMPLE 1

Ligation of the Nucleoside Diphosphate pCp to the 3' End of mRNA.

One µg of RNA was incubated in a final reaction medium of 10 µl in the presence of 5 U of T_4 phage RNA ligase in the buffer provided by the manufacturer (Gibco - BRL), 40 U of the RNase inhibitor RNasin (Promega) and, 2 µl of ^{32}pCp (Amersham #PB 10208). The incubation was performed at 37°C for 2 hours or overnight at 7-8°C.

Following modification or elimination of the 2′, 3′-cis diol at the 3′ ribose, the 2′, 3′-cis diol present at the 5′ end of the mRNA may be oxidized using reagents such as $NaBH_4$, $NaBH_3CN$, or sodium periodate, thereby converting the 2′, 3′-cis diol to a dialdehyde.

Example 2 describes the oxidation of the 2′, 3′-cis diol at the 5′ end of the mRNA with sodium periodate.

EXAMPLE 2

Oxidation of 2′, 3′-cis diol at the 5′ End of the mRNA with Sodium Periodate

0.1 OD unit of either a capped oligoribonucleotide of 47 nucleotides (including the cap) or an uncapped oligoribonucleotide of 46 nucleotides were treated as follows. The oligoribonucleotides were produced by *in vitro* transcription using the transcription kit "AmpliScribe T7" (Epicentre Technologies). As indicated below, the DNA template for the RNA transcript contained a single cytosine. To synthesize the uncapped RNA, all four NTPs were included in the *in vitro* transcription reaction. To obtain the capped RNA, GTP was replaced by an analogue of the cap, m7G(5′)ppp(5′)G. This compound, recognized by the polymerase, was incorporated into the 5′ end of the nascent transcript during the initiation of transcription but was not incorporated during the extension step. Consequently, the resulting RNA contained a cap at its 5′ end. The sequences of the oligoribonucleotides produced by the *in vitro* transcription reaction were:

+Cap:

5′m7GpppGCAUCCUACUCCCAUCCAAUUCCACCCUAACUCCUCCCAUCUCCAC-3′ (SEQ ID NO:1)

-Cap:

5′-pppGCAUCCUACUCCCAUCCAAUUCCACCCUAACUCCUCCCAUCUCCAC-3′ (SEQ ID NO:2)

The oligoribonucleotides were dissolved in 9 μl of acetate buffer (0.1 M sodium acetate, pH 5.2) and 3 μl of freshly prepared 0.1 M sodium periodate solution. The mixture was incubated for 1 hour in the dark at 4°C or room temperature. Thereafter, the reaction was stopped by adding 4 μl of 10% ethylene glycol. The product was ethanol precipitated, resuspended in at least 10 μl of water or appropriate buffer and dialyzed against water.

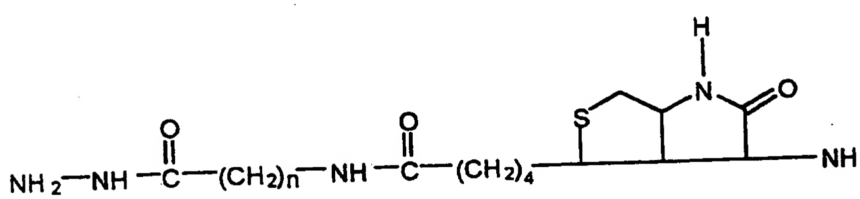
The resulting aldehyde groups may then be coupled to molecules having a reactive amine group, such as hydrazine, carbazide, thiocarbazide or semicarbazide groups, in order to facilitate enrichment of the 5′ ends of the mRNAs. Molecules having reactive amine groups

which are suitable for use in selecting mRNAs having intact 5′ ends include avidin, proteins, antibodies, vitamins, ligands capable of specifically binding to receptor molecules, or oligonucleotides. Example 3 below describes the coupling of the resulting dialdehyde to biotin.

EXAMPLE 3

Coupling of the Dialdehyde at the 5' End of Transcripts with Biotin

The oxidation product obtained in Example 2 was dissolved in 50 µl of sodium acetate at a pH between 5 and 5.2 and 50 µl of freshly prepared 0.02 M solution of biotin hydrazide in a methoxyethanol/water mixture (1:1) of formula:



In the compound used in these experiments, n=5. However, it will be appreciated that other commercially available hydrazides may also be used, such as molecules of the above formula in which n varies from 0 to 5. The mixture was then incubated for 2 hours at 37°C, precipitated with ethanol and dialyzed against distilled water. Example 4 demonstrates the specificity of the biotinylation reaction.

EXAMPLE 4

Specificity of Biotinylation of Capped Transcripts

The specificity of the biotinylation for capped mRNAs was evaluated by gel electrophoresis of the following samples:

Sample 1. The 46 nucleotide uncapped *in vitro* transcript prepared as in Example 2 and labeled with ^{32}pCp as described in Example 1.

Sample 2. The 46 nucleotide uncapped *in vitro* transcript prepared as in Example 2, labeled with ^{32}pCp as described in Example 1, treated with the oxidation reaction of Example 2, and subjected to the biotinylation conditions of Example 3.

Sample 3. The 47 nucleotide capped *in vitro* transcript prepared as in Example 2 and labeled with ^{32}pCp as described in Example 1.

Sample 4. The 47 nucleotide capped *in vitro* transcript prepared as in Example 2, labeled with ^{32}pCp as described in Example 1, treated with the oxidation reaction of Example 2, and subjected to the biotinylation conditions of Example 3.

Samples 1 and 2 had identical migration rates, demonstrating that the uncapped RNAs were not oxidized and biotinylated. Sample 3 migrated more slowly than Samples 1 and 2, while Sample 4 exhibited the slowest migration. The difference in migration of the RNAs in Samples 3 and 4 demonstrates that the capped RNAs were specifically biotinylated.

In some cases, mRNAs having intact 5′ ends may be enriched by binding the molecule containing a reactive amine group to a suitable solid phase substrate such as the inside of the vessel containing the mRNAs, magnetic beads, chromatography matrices, or nylon or nitrocellulose membranes. For example, where the molecule having a reactive amine group is biotin, the solid phase substrate may be coupled to avidin or streptavidin. Alternatively, where the molecule having the reactive amine group is an antibody or receptor ligand, the solid phase substrate may be coupled to the cognate antigen or receptor. Finally, where the molecule having a reactive amine group comprises an oligonucleotide, the solid phase substrate may comprise a complementary oligonucleotide.

The mRNAs having intact 5′ ends may be released from the solid phase following the enrichment procedure. For example, where the dialdehyde is coupled to biotin hydrazide and the solid phase comprises streptavidin, the mRNAs may be released from the solid phase by simply heating to 95 degrees Celsius in 2% SDS. In some methods, the molecule having a reactive amine group may also be cleaved from the mRNAs having intact 5′ ends following enrichment. Example 5 describes the capture of biotinylated mRNAs with streptavidin coated beads and the release of the biotinylated mRNAs from the beads following enrichment.

EXAMPLE 5

Capture and Release of Biotinylated mRNAs Using Streptavidin Coated Beads

The streptavidin coated magnetic beads were prepared according to the manufacturer's instructions (CPG Inc., USA). The biotinylated mRNAs were added to a hybridization buffer (1.5 M NaCl, pH 5 - 6). After incubating for 30 minutes, the unbound and nonbiotinylated material was removed. The beads were then washed several times in water with 1% SDS. The beads thus obtained were incubated for 15 minutes at 95°C in water containing 2% SDS.

Example 6 demonstrates the efficiency with which biotinylated mRNAs were recovered from the streptavidin coated beads.

EXAMPLE 6

Efficiency of Recovery of Biotinylated mRNAs

The efficiency of the recovery procedure was evaluated as follows. Capped RNAs were labeled with ^{32}pCp, oxidized, biotinylated and bound to streptavidin coated beads as described above. Subsequently, the bound RNAs were incubated for 5, 15 or 30 minutes at 95°C in the presence of 2% SDS.

The products of the reaction were analyzed by electrophoresis on 12% polyacrylamide gels under denaturing conditions (7 M urea). The gels were subjected to autoradiography. During this manipulation, the hydrazone bonds were not reduced.

Increasing amounts of nucleic acids were recovered as incubation times in 2% SDS increased, demonstrating that biotinylated mRNAs were efficiently recovered.

In an alternative method for obtaining mRNAs having intact 5′ ends, an oligonucleotide which has been derivatized to contain a reactive amine group is specifically coupled to mRNAs having an intact cap. Preferably, the 3′ end of the mRNA is blocked prior to the step in which the aldehyde groups are joined to the derivatized oligonucleotide, as described above, so as to prevent the derivatized oligonucleotide from being joined to the 3′ end of the mRNA. For example, pCp may be attached to the 3′ end of the mRNA using T4 RNA ligase as described in example 1. However, as discussed above, blocking the 3′ end of

the mRNA is an optional step. Derivatized oligonucleotides may be prepared as described in Example 7.

EXAMPLE 7

Derivatization of Oligonucleotides

An oligonucleotide phosphorylated at its 3′ end was converted to a 3′ hydrazide in 3′ by treatment with an aqueous solution of hydrazine or of dihydrazide of the formula $H_2N(R1)NH_2$ at about 1 to 3 M, and at pH 4.5 at a temperature of 8°C overnight. This incubation was performed in the presence of a carbodiimide type agent soluble in water such as 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide at a final concentration of 0.3 M.

The derivatized oligonucleotide was then separated from the other agents and products using a standard technique for isolating oligonucleotides.

As discussed above, the mRNAs to be enriched may be treated to eliminate the 3′ OH groups which may be present thereon. This may be accomplished by enzymatic ligation of sequences lacking a 3′ OH, such as pCp, as described in Example 1. Alternatively, the 3′ OH groups may be eliminated by alkaline hydrolysis as described in Example 8 below.

EXAMPLE 8

Elimination of 3' OH Groups of mRNA Using Alkaline Hydrolysis

In a total volume of 100 µl of 0.1 N sodium hydroxide, 1.5 µg mRNA is incubated for 40 to 60 minutes at 4°C. The solution is neutralized with acetic acid and precipitated with ethanol.

Following the optional elimination of the 3′ OH groups, the diol groups at the 5′ ends of the mRNAs are oxidized as described below in Example 9.

EXAMPLE 9

Oxidation of Diols of mRNA

Up to 1 OD unit of RNA was dissolved in 9 µl of buffer (0.1 M sodium acetate, pH 6-7) or water and 3 µl of freshly prepared 0.1 M sodium periodate solution. The reaction was incubated for 1 h in the dark at 4°C or room temperature. Following the incubation, the reaction was stopped by adding 4 µl of 10% ethylene glycol. Thereafter the mixture was

incubated at room temperature for 15 minutes. After ethanol precipitation, the product was resuspended in at least 10 μl of water or appropriate buffer and dialyzed against water.

Following oxidation of the diol groups at the 5′ ends of the mRNAs, the derivatized oligonucleotide was joined to the resulting aldehydes as described in Example 10.

EXAMPLE 10

Ligature of Aldehydes of mRNA to Derivatized Oligonucleotides

The oxidized mRNA was dissolved in an acidic medium such as 50 μl of sodium acetate pH 4-6. Fifty μl of a solution of the derivatized oligonucleotide were added in order to obtain an mRNA:derivatized oligonucleotide ratio of 1:20. The mixture was reduced with a borohydride and incubated for 2 h at 37°C or overnight (14 h) at 10°C. The mixture was then ethanol precipitated, resuspended in 10 μl or more of water or appropriate buffer and dialyzed against distilled water. If desired, the resulting product may be analyzed using acrylamide gel electrophoresis, HPLC analysis, or other conventional techniques.

Following the attachment of the derivatized oligonucleotide to the mRNAs, a reverse transcription reaction may be performed as described in Example 11 below.

EXAMPLE 11

Reverse Transcription of mRNAs Ligatured to Derivatized Oligonucleotides

An oligodeoxyribonucleotide was derivatized as follows. Three OD units of an oligodeoxyribonucleotide of sequence 5'ATCAAGAATTCGCACGAGACCATTA3' (SEQ ID NO:3) having 5′-OH and 3′-P ends were dissolved in 70 μl of a 1.5 M hydroxybenzotriazole solution, pH 5.3, prepared in dimethylformamide/water (75:25) containing 2 μg of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide. The mixture was incubated for 2 h 30 min at 22°C and then precipitated twice in $LiClO_4$/acetone. The pellet was resuspended in 200 μl of 0.25 M hydrazine and incubated at 8°C from 3 to 14 h. Following the hydrazine reaction, the mixture was precipitated twice in $LiClO_4$/acetone.

The messenger RNAs to be reverse transcribed were extracted from blocks of placenta having sides of 2 cm which had been stored at -80°C. The total RNA was extracted

using conventional acidic phenol techniques. Oligo-dT chromatography was used to purify the mRNAs. The integrity of the mRNAs was checked by Northern-blotting.

The diol groups on 7 µg of the placental mRNAs were oxidized as described above in Example 9. The derivatized oligonucleotide was joined to the mRNAs as described in Example 10 above except that the precipitation step was replaced by an exclusion chromatography step to remove derivatized oligodeoxyribonucleotides which were not joined to mRNAs. Exclusion chromatography was performed as follows:

Ten ml of Ultrogel AcA34 (BioSepra#230151) gel, a mix of agarose and acrylamide, were equilibrated in 50 ml of a solution of 10 mM Tris pH 8.0, 300 mM NaCl, 1 mM EDTA, and 0.05% SDS. The mixture was allowed to sediment. The supernatant was eliminated and the gel was resuspended in 50 ml of buffer. This procedure was repeated 2 or 3 times.

A glass bead (diameter 3 mm) was introduced into a 2 ml disposable pipette (length 25 cm). The pipette was filled with the gel suspension until the height of the gel stabilized at 1 cm from the top of the pipette. The column was then equilibrated with 20 ml of equilibration buffer (10 mM Tris HCl pH 7.4, 20 mM NaCl).

Ten µl of the mRNA which had reacted with the derivatized oligonucleotide were mixed in 39 µl of 10 mM urea and 2 µl of blue-glycerol buffer, which had been prepared by dissolving 5 mg of bromophenol blue in 60% glycerol (v/v), and passing the mixture through a 0.45 µm diameter filter.

The column was then loaded with the mRNAs coupled to the oligonucleotide. As soon as the sample had penetrated, equilibration buffer was added. Hundred µl fractions were then collected. Derivatized oligonucleotide which had not been attached to mRNA appeared in fraction 16 and later fractions. Thus, fractions 3 to 15 were combined and precipitated with ethanol.

To determine whether the derivatized oligonucleotide was actually linked to mRNA, one tenth of the combined fractions were spotted twice on a nylon membrane and hybridized to a radioactive probe using conventional techniques. The ^{32}P labeled probe used in these hybridizations was an oligodeoxyribonucleotide of sequence 5'TAATGGTCTCGTGCGAATTCTTGAT3' (SEQ ID NO:4) anticomplementary to the derivatized oligonucleotide. A signal observed after autoradiography, indicated that the derivatized oligonucleotide had been truly joined to the mRNA.

The remaining nine tenth of the mRNAs which had reacted with the derivatized oligonucleotide was reverse transcribed as follows. A reverse transcription reaction was carried out with reverse transcriptase following the manufacturer's instructions and 50 pmol of nonamers with random sequence as primers.

To ensure that reverse transcription had been carried out through the cap structure, two types of experiments were performed.

In the first approach, after elimination of RNA of the cDNA:RNA heteroduplexes obtained from the reverse transcription reaction by an alkaline hydrolysis, a portion of the resulting single stranded cDNAs was spotted on a positively charged membrane and hybridized, using conventional methods, to a ^{32}P labeled probe having a sequence identical to that of the derivatized oligonucleotide. Control spots containing, 1 pmol, 100 fmol, 50 fmol, 10 fmol and 1 fmol of a control oligodeoxyribonucleotide of sequence identical to that of the derivatized oligonucleotide were included. The signal observed in the spots containing the cDNA indicated that approximately 15 fmol of the derivatized oligonucleotide had been reverse transcribed. These results demonstrate that the reverse transcription can be performed through the cap and, in particular, that reverse transcriptase crosses the 5'-P-P-P-5' bond of the cap of eukaryotic messenger RNAs.

In the second type of experiment, the single stranded cDNAs obtained from the above first strand synthesis were used as template for PCR reactions. Two types of reactions were carried out. First, specific amplification of the mRNAs for alpha globin, dehydrogenase, pp15 and elongation factor E4 were carried out using the following pairs of oligodeoxyribonucleotide primers.

alpha-globin

GLO-S: 5'CCG ACA AGA CCA ACG TCA AGG CCG C3' (SEQ ID NO:5)

GLO-As: 5'TCA CCA GCA GGC AGT GGC TTA GGA G 3' (SEQ ID NO:6)

dehydrogenase

3 DH-S: 5'AGT GAT TCC TGC TAC TTT GGA TGG C3' (SEQ ID NO:7)

3 DH-As: 5'GCT TGG TCT TGT TCT GGA GTT TAG A3' (SEQ ID NO:8)

pp15

PP15-S: 5'TCC AGA ATG GGA GAC AAG CCA ATT T3' (SEQ ID NO:9)

PP15-As: 5'AGG GAG GAG GAA ACA GCG TGA GTC C3' (SEQ ID NO:10)

Elongation factor E4

EFA1-S: 5'ATG GGA AAG GAA AAG ACT CAT ATC A3' (SEQ ID NO:11)

EF1A-As: 5'AGC AGC AAC AAT CAG GAC AGC ACA G3' (SEQ ID NO:12)

Second, non specific amplifications were also carried out with the antisense oligodeoxyribonucleotides of the pairs described above and with a primer derived from the sequence of the derivatized oligodeoxyribonucleotide (5'ATCAAGAATTCGCACGAGACCATTA3') (SEQ ID NO:13).

One twentieth of the following RT-PCR product samples were run on a 1.5% agarose gel and stained with ethidium bromide.

Sample 1: The products of a PCR reaction using the globin primers of SEQ ID NOs 5 and 6 in the presence of cDNA.

Sample 2: The products of a PCR reaction using the globin primers of SEQ ID NOs 5 and 6 in the absence of added cDNA.

Sample 3: The products of a PCR reaction using the dehydrogenase primers of SEQ ID NOs 7 and 8 in the presence of cDNA.

Sample 4: The products of a PCR reaction using the dehydrogenase primers of SEQ ID NOs 7 and 8 in the absence of added cDNA.

Sample 5: The products of a PCR reaction using the pp15 primers of SEQ ID NOs 9 and 10 in the presence of cDNA.

Sample 6: The products of a PCR reaction using the pp15 primers of SEQ ID NOs 9 and 10 in the absence of added cDNA.

Sample 7: The products of a PCR reaction using the EIF4 primers of SEQ ID NOs 11 and 12 in the presence of added cDNA.

Sample 8: The products of a PCR reaction using the EIF4 primers of SEQ ID NOs 11 and 12 in the absence of added cDNA.

A band of the size expected for the PCR product was observed only in samples 1, 3, 5 and 7, thus indicating the presence of the corresponding sequence in the cDNA population.

PCR reactions were also carried out with the antisense oligonucleotides of the globin and dehydrogenase primers (SEQ ID NOs 6 and 8) and an oligonucleotide whose sequence corresponds to that of the derivatized oligonucleotide. The presence of PCR products of the expected size in the samples equivalent to above samples 1 and 3 indicated that the derivatized oligonucleotide had been linked to mRNA.

The above examples summarize the chemical procedure for enriching mRNAs for those having intact 5′ ends as illustrated in Figure 1. Further detail regarding the chemical approaches for obtaining such mRNAs are disclosed in International Application No. WO96/34981, published November 7, 1996, which is incorporated herein by reference. Strategies based on the above chemical modifications to the 5′ cap structure may be utilized to generate cDNAs selected to include the 5′ ends of the mRNAs from which they derived. In one version of such procedures, the 5′ ends of the mRNAs are modified as described above. Thereafter, a reverse transcription reaction is conducted to extend a primer complementary to the 5′ end of the mRNA. Single stranded RNAs are eliminated to obtain a population of cDNA/mRNA heteroduplexes in which the mRNA includes an intact 5′ end. The resulting heteroduplexes may be captured on a solid phase coated with a molecule capable of interacting with the molecule used to derivatize the 5′ end of the mRNA. Thereafter, the strands of the heteroduplexes are separated to recover single stranded first cDNA strands which include the 5′ end of the mRNA. Second strand cDNA synthesis may then proceed using conventional techniques. For example, the procedures disclosed in WO 96/34981 or in Carninci. *et al., Genomics* 37:327-336, 1996, the disclosures of which are incorporated herein by reference, may be employed to select cDNAs which include the sequence derived from the 5′ end of the coding sequence of the mRNA.

Following ligation of the oligonucleotide tag to the 5′ cap of the mRNA, a reverse transcription reaction is conducted to extend a primer complementary to the mRNA to the 5′ end of the mRNA. Following elimination of the RNA component of the resulting heteroduplex using standard techniques, second strand cDNA synthesis is conducted with a primer complementary to the oligonucleotide tag.

2. Enzymatic Methods for Obtaining mRNAs having Intact 5′ Ends

Other techniques for selecting cDNAs extending to the 5′ end of the mRNA from which they are derived are fully enzymatic. Some versions of these techniques are disclosed in Dumas Milne Edwards J.B. (Doctoral Thesis of Paris VI University, Le clonage des ADNc complets: difficultes et perspectives nouvelles. Apports pour l'etude de la regulation de l'expression de la tryptophane hydroxylase de rat, 20 Dec. 1993), EP0 625572 and Kato *et al.*, *Gene* **150**:243-250, 1994, the disclosures of which are incorporated herein by reference.

Briefly, in such approaches, isolated mRNA is treated with alkaline phosphatase to remove the phosphate groups present on the 5′ ends of uncapped incomplete mRNAs. Following this procedure, the cap present on full length mRNAs is enzymatically removed with a decapping enzyme such as T4 polynucleotide kinase or tobacco acid pyrophosphatase. An oligonucleotide, which may be either a DNA oligonucleotide or a DNA-RNA hybrid oligonucleotide having RNA at its 3′ end, is then ligated to the phosphate present at the 5′ end of the decapped mRNA using T4 RNA ligase. The oligonucleotide may include a restriction site to facilitate cloning of the cDNAs following their synthesis. Example 12 below describes one enzymatic method based on the doctoral thesis of Dumas.

EXAMPLE 12

Enzymatic Approach for Obtaining 5′ ESTs

Twenty micrograms of PolyA+ RNA were dephosphorylated using Calf Intestinal Phosphatase (Biolabs). After a phenol chloroform extraction, the cap structure of mRNA was hydrolysed using the Tobacco Acid Pyrophosphatase (purified as described by Shinshi *et al.*., *Biochemistry* 15: 2185-2190, 1976) and a hemi 5′DNA/RNA-3′ oligonucleotide having an unphosphorylated 5′ end, a stretch of adenosine ribophosphate at the 3′ end, and an EcoRI site near the 5′ end was ligated to the 5′P ends of mRNA using the T4 RNA ligase (Biolabs). Oligonucleotides suitable for use in this procedure are preferably 30 to 50 bases in length. Oligonucleotides having an unphosphorylated 5′ end may be synthesized by adding a fluorochrome at the 5′ end. The inclusion of a stretch of adenosine ribophosphates at the 3′ end of the oligonucleotide increases ligation efficiency. It will be appreciated that the oligonucleotide may contain cloning sites other than EcoRI.

Following ligation of the oligonucleotide to the phosphate present at the 5′ end of the decapped mRNA, first and second strand cDNA synthesis is carried out using conventional methods or those specified in EP0 625,572 and Kato *et al. supra*, and Dumas Milne Edwards, *supra*, the disclosures of which are incorporated herein by reference. The resulting cDNA may then be ligated into vectors such as those disclosed in Kato *et al.*, *supra* or other nucleic acid vectors known to those skilled in the art using techniques such as those described in Sambrook *et al.*, Molecular Cloning: A Laboratory Manual 2d Ed., Cold Spring Harbor Laboratory Press, 1989, the disclosure of which is incorporated herein by reference.

II. Obtention and Characterization of the 5′ ESTs of the Present Invention

The 5' ESTs of the present invention were obtained using the aforementioned chemical and enzymatic approaches for enriching mRNAs for those having intact 5′ ends as decribed below.

1. Obtention of 5' ESTS Using mRNAs with Intact 5' Ends

First, mRNAs were prepared as described in Example 13 below.

EXAMPLE 13

Preparation of mRNA With Intact 5' Ends

Total human RNAs or polyA$^+$ RNAs derived from 29 different tissues were respectively purchased from LABIMO and CLONTECH and used to generate 44 cDNA libraries as follows. The purchased RNA had been isolated from cells or tissues using acid guanidium thiocyanate-phenol-chloroform extraction (Chomczyniski and Sacchi, *Analytical Biochemistry* 162:156-159, 1987). PolyA$^+$ RNA was isolated from total RNA (LABIMO) by two passes of oligo dT chromatography, as described by Aviv and Leder, *Proc. Natl. Acad. Sci. USA* 69:1408-1412, 1972 in order to eliminate ribosomal RNA.

The quality and the integrity of the polyA+ RNAs were checked. Northern blots hybridized with a globin probe were used to confirm that the mRNAs were not degraded. Contamination of the polyA$^+$ mRNAs by ribosomal sequences was checked using Northern blots and a probe derived from the sequence of the 28S rRNA. Preparations of mRNAs with

less than 5% of rRNAs were used in library construction. To avoid constructing libraries with RNAs contaminated by exogenous sequences (prokaryotic or fungal), the presence of bacterial 16S ribosomal sequences or of two highly expressed fungal mRNAs was examined using PCR.

Following preparation of the mRNAs, the above described chemical and/or the enzymatic procedures for enriching mRNAs for thoses having intact 5′ ends were employed to obtain 5′ ESTs from various tissues. In both approaches, an oligonucleotide tag was attached to the 5′ ends of the mRNAs. The oligonucleotide tag had an EcoRI site therein to facilitate later cloning procedures. To facilitate the processing of single stranded and double stranded cDNA obtained in the construction of the librairies, the same nucleotidic sequence was used to design the ligated oligonucleotide in both chemical and enzymatic approaches. Nevertheless, in the chemical procedure, the tag used was an oligodeoxyribonucleotide which was linked to the cap of the mRNA whereas in the enzymatic ligation, the tag was a chimeric hemi 5'DNA/RNA3' oligonucleotide which was ligated to the 5' end of decapped mRNA as described in example 12.

Following attachment of the oligonucleotide tag to the mRNA by either the chemical or enzymatic methods, the integrity of the mRNA was examined by performing a Northern blot with 200 to 500 ng of mRNA using a probe complementary to the oligonucleotide tag before performing the first strand synthesis as described in example 14.

EXAMPLE 14

<u>cDNA Synthesis Using mRNA Templates Having Intact 5′ Ends</u>

For the mRNAs joined to oligonucleotide tags using both the chemical and enzymatic methods, first strand cDNA synthesis was performed using the Superscript II (Gibco BRL) or the Rnase H Minus M-MLV (Promega) reverse transcriptase with random nonamers as primers. In order to protect internal EcoRI sites in the cDNA from digestion at later steps in the procedure, methylated dCTP was used for first strand synthesis. After removal of RNA by an alkaline hydrolysis, the first strand of cDNA was precipitated using isopropanol in order to eliminate residual primers.

For both the chemical and the enzymatic methods, the second strand of the cDNA was synthesized with a Klenow fragment using a primer corresponding to the 5′ end of the

ligated oligonucleotide described in Example 12. Preferably, the primer is 20-25 bases in length. Methylated dCTP was also used for second strand synthesis in order to protect internal EcoRI sites in the cDNA from digestion during the cloning process.

Following cDNA synthesis, the cDNAs were cloned into pBlueScript as described in Example 15 below.

EXAMPLE 15

Cloning of cDNAsderived from mRNA with intact 5' ends into BlueScript

Following second strand synthesis, the ends of the cDNA were blunted with T4 DNA polymerase (Biolabs) and the cDNA was digested with EcoRI. Since methylated dCTP was used during cDNA synthesis, the EcoRI site present in the tag was the only hemi-methylated site, hence the only site susceptible to EcoRI digestion. The cDNA was then size fractionated using exclusion chromatography (AcA, Biosepra) and fractions corresponding to cDNAs of more than 150 bp were pooled and ethanol precipitated. The cDNA was directionally cloned into the SmaI and EcoRI ends of the phagemid pBlueScript vector (Stratagene). The ligation mixture was electroporated into bacteria and propagated under appropriate antibiotic selection.

Clones containing the oligonucleotide tag attached were then selected as described in Example 16 below.

EXAMPLE 16

Selection of Clones Having the Oligonucleotide Tag Attached Thereto

The plasmid DNAs containing 5′ EST libraries made as described above were purified (Qiagen). A positive selection of the tagged clones was performed as follows. Briefly, in this selection procedure, the plasmid DNA was converted to single stranded DNA using gene II endonuclease of the phage F1 in combination with an exonuclease (Chang *et al.*, *Gene* 127:95-8, 1993) such as exonuclease III or T7 gene 6 exonuclease. The resulting single stranded DNA was then purified using paramagnetic beads as described by Fry *et al.*, *Biotechniques*, **13**: 124-131, 1992. In this procedure, the single stranded DNA was hybridized with a biotinylated oligonucleotide having a sequence corresponding to the 3′ end of the oligonucleotide described in Example 13. Preferably, the primer has a length of 20-25

bases. Clones including a sequence complementary to the biotinylated oligonucleotide were captured by incubation with streptavidin coated magnetic beads followed by magnetic selection. After capture of the positive clones, the plasmid DNA was released from the magnetic beads and converted into double stranded DNA using a DNA polymerase such as the ThermoSequenase obtained from Amersham Pharmacia Biotech. Alternatively, protocoles such as the one described in the Gene Trapper kit available from Gibco BRL may be used. The double stranded DNA was then electroporated into bacteria. The percentage of positive clones having the 5' tag oligonucleotide was estimated to typically rank between 90 and 98% using dot blot analysis.

Following electroporation, the libraries were ordered in 384-microtiter plates (MTP). A copy of the MTP was stored for future needs. Then the libraries were transferred into 96 MTP and sequenced as described below.

EXAMPLE 17

Sequencing of Inserts in Selected Clones

Plasmid inserts were first amplified by PCR on PE 9600 thermocyclers (Perkin-Elmer, Applied Biosystems Division, Foster City, CA), using standard SETA-A and SETA-B primers (Genset SA), AmpliTaqGold (Perkin-Elmer), dNTPs (Boehringer), buffer and cycling conditions as recommended by the Perkin-Elmer Corporation.

PCR products were then sequenced using automatic ABI Prism 377 sequencers (Perkin Elmer). Sequencing reactions were performed using PE 9600 thermocyclers with standard dye-primer chemistry and ThermoSequenase (Amersham Pharmacia Biotech). The primers used were either T7 or 21M13 (available from Genset SA) as appropriate. The primers were labeled with the JOE, FAM, ROX and TAMRA dyes. The dNTPs and ddNTPs used in the sequencing reactions were purchased from Boehringer. Sequencing buffer, reagent concentrations and cycling conditions were as recommended by Amersham.

Following the sequencing reaction, the samples were precipitated with ethanol, resuspended in formamide loading buffer, and loaded on a standard 4% acrylamide gel. Electrophoresis was performed for 2.5 hours at 3000V on an ABI 377 sequencer, and the sequence data were collected and analyzed using the ABI Prism DNA Sequencing Analysis Software, version 2.1.2.

2. Computer analysis of the Obtained 5' ESTs: Construction of NetGene and SignalTag databases

The sequence data from the 44 cDNA libraries made as described above were transferred to a proprietary database, where quality control and validation steps were performed. A proprietary base-caller, working using a Unix system, automatically flagged suspect peaks, taking into account the shape of the peaks, the inter-peak resolution, and the noise level. The proprietary base-caller also performed an automatic trimming. Any stretch of 25 or fewer bases having more than 4 suspect peaks was considered unreliable and was discarded. Sequences corresponding to cloning vector or ligation oligonucleotides were automatically removed from the EST sequences. However, the resulting EST sequences may contain 1 to 5 bases belonging to the above mentioned sequences at their 5' end. If needed, these can easily be removed on a case to case basis.

Following sequencing as described above, the sequences of the 5' ESTs were entered in NetGene™, a proprietary database called for storage and manipulation as described below. It will be appreciated by those skilled in the art that the data could be stored and manipulated on any medium which can be read and accessed by a computer. Computer readable media include magnetically, optically, or electronically readable media. For example, the computer readable media may be a hard disc, a floppy disc, a magnetic tape, CD-ROM, RAM, or ROM as well as other types of other media known to those skilled in the art.

In addition, the sequence data may be stored and manipulated in a variety of data processor programs in a diversity of formats. For instance, the sequence data may be stored as text in a word processing file, such as Microsoft WORD or WORDPERFECT or as an ASCII file in a variety of database programs familiar to those of skill in the art, such as DB2, SYBASE, or ORACLE.

The computer readable media on which the sequence information is stored may be in a personal computer, a network, a server or other computer systems known to those skilled in the art. The computer or other system preferably includes the storage media described above, and a processor for accessing and manipulating the sequence data. Once the sequence data has been stored, it may be manipulated and searched to locate those stored sequences which contain a desired nucleic acid sequence or which encode a protein having a particular functional domain. For example, the stored sequence information may be compared to other

known sequences to identify homologies, motifs implicated in biological function, or structural motifs.

Programs which may be used to search or compare the stored sequences include the MacPattern (EMBL), BLAST, and BLAST2 program series (NCBI), basic local alignment search tool programs for nucleotide (BLASTN) and peptide (BLASTX) comparisons (Altschul *et al*, *J. Mol. Biol.* 215: 403, 1990) and FASTA (Pearson and Lipman, *Proc. Natl. Acad. Sci. USA* 85: 2444, 1988). The BLAST programs then extend the alignments on the basis of defined match and mismatch criteria.

Motifs which may be detected using the above programs and those described in Example 28 include sequences encoding leucine zippers, helix-turn-helix motifs, glycosylation sites, ubiquitination sites, alpha helices, and beta sheets, signal sequences encoding signal peptides which direct the secretion of the encoded proteins, sequences implicated in transcription regulation such as homeoboxes, acidic stretches, enzymatic active sites, substrate binding sites, and enzymatic cleavage sites.

Before searching the cDNAs in the NetGene™ database for sequence motifs of interest, cDNAs derived from mRNAs which were not of interest were identified and eliminated from further consideration as described in Example 18 below.

EXAMPLE 18

Elimination of Undesired Sequences from Further Consideration

5′ ESTs in the NetGene™ database which were derived from undesired sequences such as transfer RNAs, ribosomal RNAs, mitochondrial RNAs, prokaryotic RNAs, fungal RNAs, Alu sequences, L1 sequences, or repeat sequences were identified using the FASTA and BLASTN programs with the parameters listed in Table I.

To eliminate 5′ ESTs encoding tRNAs from further consideration, the 5′ EST sequences were compared to the sequences of 1190 known tRNAs obtained from EMBL release 38, of which 100 were human. The comparison was performed using FASTA on both strands of the 5′ ESTs. Sequences having more than 80% homology over more than 60 nucleotides were identified as tRNA. Of the 144,341 sequences screened, 26 were identified as tRNAs and eliminated from further consideration.

To eliminate 5′ ESTs encoding rRNAs from further consideration, the 5′ EST sequences were compared to the sequences of 2497 known rRNAs obtained from EMBL release 38, of which 73 were human. The comparison was performed using BLASTN on both strands of the 5′ ESTs with the parameter S=108. Sequences having more than 80% homology over stretches longer than 40 nucleotides were identified as rRNAs. Of the 144,341 sequences screened, 3,312 were identified as rRNAs and eliminated from further consideration.

To eliminate 5′ ESTs encoding mtRNAs from further consideration, the 5′ EST sequences were compared to the sequences of the two known mitochondrial genomes for which the entire genomic sequences are available and all sequences transcribed from these mitochondrial genomes including tRNAs, rRNAs, and mRNAs for a total of 38 sequences. The comparison was performed using BLASTN on both strands of the 5′ ESTs with the parameter S=108. Sequences having more than 80% homology over stretches longer than 40 nucleotides were identified as mtRNAs. Of the 144,341 sequences screened, 6,110 were identified as mtRNAs and eliminated from further consideration.

Sequences which might have resulted from exogenous contaminants were eliminated from further consideration by comparing the 5′ EST sequences to release 46 of the EMBL bacterial and fungal divisions using BLASTN with the parameter S=144. All sequences having more than 90% homology over at least 40 nucleotides were identified as exogenous contaminants. Of the 42 cDNA libraries examined, the average percentages of prokaryotic and fungal sequences contained therein were 0.2% and 0.5% respectively. Among these sequences, only one could be identified as a sequence specific to fungi. The others were either fungal or prokaryotic sequences having homologies with vertebrate sequences or including repeat sequences which had not been masked during the electronic comparison.

In addition, the 5′ ESTs were compared to 6093 Alu sequences and 1115 L1 sequences to mask 5′ ESTs containing such repeat sequences. 5′ ESTs including THE and MER repeats, SSTR sequences or satellite, micro-satellite, or telomeric repeats were also eliminated from further consideration. On average, 11.5% of the sequences in the libraries contained repeat sequences. Of this 11.5%, 7% contained Alu repeats, 3.3% contained L1 repeats and the remaining 1.2% were derived from the other screened types of repetitive sequences. These percentages are consistent with those found in cDNA libraries prepared by

other groups. For example, the cDNA libraries of Adams *et al.* contained between 0% and 7.4% Alu repeats depending on the source of the RNA which was used to prepare the cDNA library (Adams *et al.*, *Nature* 377:174, 1996).

The sequences of those 5′ ESTs remaining after the elimination of undesirable sequences were compared with the sequences of known human mRNAs to determine the accuracy of the sequencing procedures described above.

EXAMPLE 19

Measurement of Sequencing Accuracy by Comparison to Known Sequences

To further determine the accuracy of the sequencing procedure described above, the sequences of 5′ ESTs derived from known sequences were identified and compared to the original known sequences. First, a FASTA analysis with overhangs shorter than 5 bp on both ends was conducted on the 5′ ESTs to identify those matching an entry in the public human mRNA database. The 6655 5′ ESTs which matched a known human mRNA were then realigned with their cognate mRNA and dynamic programming was used to include substitutions, insertions, and deletions in the list of "errors" which would be recognized. Errors occurring in the last 10 bases of the 5′ EST sequences were ignored to avoid the inclusion of spurious cloning sites in the analysis of sequencing accuracy.

This analysis revealed that the sequences incorporated in the NetGene™ database had an accuracy of more than 99.5%.

To determine the efficiency with which the above selection procedures select cDNAs which include the 5′ ends of their corresponding mRNAs, the following analysis was performed.

EXAMPLE 20

Determination of Efficiency of 5′ EST Selection

To determine the efficiency at which the above selection procedures isolated 5′ ESTs which included sequences close to the 5′ end of the mRNAs from which they derived, the sequences of the ends of the 5′ ESTs derived from the elongation factor 1 subunit α and

ferritin heavy chain genes were compared to the known cDNA sequences of these genes. Since the transcription start sites of both genes are well characterized, they may be used to determine the percentage of derived 5′ ESTs which included the authentic transcription start sites.

For both genes, more than 95% of the obtained 5' ESTs actually included sequences close to or upstream of the 5′ end of the corresponding mRNAs.

To extend the analysis of the reliability of the procedures for isolating 5′ ESTs from ESTs in the NetGene™ database, a similar analysis was conducted using a database composed of human mRNA sequences extracted from GenBank database release 97 for comparison. The 5' ends of more than 85% of 5' ESTs derived from mRNAs included in the GeneBank database were located close to the 5′ ends of the known sequence. As some of the mRNA sequences available in the GenBank database are deduced from genomic sequences, a 5′ end matching with these sequences will be counted as an internal match. Thus, the method used here underestimates the yield of ESTs including the authentic 5′ ends of their corresponding mRNAs.

The EST libraries made above included multiple 5′ ESTs derived from the same mRNA. The sequences of such 5′ ESTs were compared to one another and the longest 5′ ESTs for each mRNA were identified. Overlapping cDNAs were assembled into continuous sequences (contigs). The resulting continuous sequences were then compared to public databases to gauge their similarity to known sequences, as described in Example 21 below.

EXAMPLE 21

Clustering of the 5′ ESTs and Calculation of Novelty Indices for cDNA Libraries

For each sequenced EST library, the sequences were clustered by the 5′ end. Each sequence in the library was compared to the others with BLASTN2 (direct strand, parameters S=107). ESTs with High Scoring Segment Pairs (HSPs) at least 25 bp long, having 95% identical bases and beginning closer than 10 bp from each EST 5′ end were grouped. The longest sequence found in the cluster was used as representative of the group. A global clustering between libraries was then performed leading to the definition of super-contigs.

To assess the yield of new sequences within the EST libraries, a novelty rate (NR) was defined as: NR= 100 X (Number of new unique sequences found in the library/Total number of sequences from the library). Typically, novelty rating ranged between 10% and 41% depending on the tissue from which the EST library was obtained. For most of the libraries, the random sequencing of 5′ EST libraries was pursued until the novelty rate reached 20%.

Following characterization as described above, the collection of 5′ ESTs in NetGene™ was screened to identify those 5′ ESTs bearing potential signal sequences as described in Example 22 below.

EXAMPLE 22

Identification of Potential Signal Sequences in 5′ ESTs

The 5′ ESTs in the NetGene™ database were screened to identify those having an uninterrupted open reading frame (ORF) longer than 45 nucleotides beginning with an ATG codon and extending to the end of the EST. Approximately half of the cDNA sequences in NetGene™ contained such an ORF. The ORFs of these 5′ ESTs were then searched to identify potential signal motifs using slight modifications of the procedures disclosed in Von Heijne, *Nucleic Acids Res.* 14:4683-4690, 1986, the disclosure of which is incorporated herein by reference. Those 5′ EST sequences encoding a stretch of at least 15 amino acid long with a score of at least 3.5 in the Von Heijne signal peptide identification matrix were considered to possess a signal sequence. Those 5′ ESTs which matched a known human mRNA or EST sequence and had a 5′ end more than 20 nucleotides downstream of the known 5′ end were excluded from further analysis. The remaining cDNAs having signal sequences therein were included in a database called SignalTag™.

To confirm the accuracy of the above method for identifying signal sequences, the analysis of Example 23 was performed.

EXAMPLE 23

Confirmation of Accuracy of Identification of Potential Signal Sequences in 5' ESTs

The accuracy of the above procedure for identifying signal sequences encoding signal peptides was evaluated by applying the method to the 43 amino acids located at the N terminus of all human SwissProt proteins. The computed Von Heijne score for each protein was compared with the known characterization of the protein as being a secreted protein or a non-secreted protein. In this manner, the number of non-secreted proteins having a score higher than 3.5 (false positives) and the number of secreted proteins having a score lower than 3.5 (false negatives) could be calculated.

Using the results of the above analysis, the probability that a peptide encoded by the 5' region of the mRNA is in fact a genuine signal peptide based on its Von Heijne's score was calculated based on either the assumption that 10 % of human proteins are secreted or the assumption that 20 % of human proteins are secreted. The results of this analysis are shown in Figure 2 and in table IV.

Using the above method of identification of secretory proteins, 5' ESTs of the following polypeptides known to be secreted were obtained: human glucagon, gamma interferon induced monokine precursor, secreted cyclophilin-like protein, human pleiotropin, and human biotinidase precursor. Thus, the above method successfully identified those 5' ESTs which encode a signal peptide.

To confirm that the signal peptide encoded by the 5' ESTs actually functions as a signal peptide, the signal sequences from the 5' ESTs may be cloned into a vector designed for the identification of signal peptides. Such vectors are designed to confer the ability to grow in selective medium only to host cells containing a vector with an operably linked signal sequence. For example, to confirm that a 5' EST encodes a genuine signal peptide, the signal sequence of the 5' EST may be inserted upstream and in frame with a non-secreted form of the yeast invertase gene in signal peptide selection vectors such as those described in U.S. Patent No. 5,536,637, the disclosure of which is incorporated herein by reference. Growth of host cells containing signal sequence selection vectors with the correctly inserted 5' EST signal sequence confirms that the 5' EST encodes a genuine signal peptide.

Alternatively, the presence of a signal peptide may be confirmed by cloning the extended cDNAs obtained using the ESTs into expression vectors such as pXT1 (as described below in example 30), or by constructing promoter-signal sequence-reporter gene vectors which encode fusion proteins between the signal peptide and an assayable reporter protein. After introduction of these vectors into a suitable host cell, such as COS cells or NIH 3T3 cells, the growth medium may be harvested and analyzed for the presence of the secreted protein. The medium from these cells is compared to the medium from control cells containing vectors lacking the signal sequence or extended cDNA insert to identify vectors which encode a functional signal peptide or an authentic secreted protein.

Those 5′ ESTs which encoded a signal peptide, as determined by the method of Example 22 above, were further grouped into four categories based on their homology to known sequences as described in Example 24 below.

EXAMPLE 24

Categorization of 5′ ESTs Encoding a Signal Peptide

Those 5′ ESTs having a sequence not matching any known vertebrate sequence nor any publicly available EST sequence were designated "new." Of the sequences in the SignalTag™ database, 947 of the 5′ ESTs having a Von Heijne's score of at least 3.5 fell into this category.

Those 5′ ESTs having a sequence not matching any vertebrate sequence but matching a publicly known EST were designated "EST-ext", provided that the known EST sequence was extended by at least 40 nucleotides in the 5′ direction. Of the sequences in the SignalTag™ database, 150 of the 5′ ESTs having a Von Heijne's score of at least 3.5 fell into this category.

Those ESTs not matching any vertebrate sequence but matching a publicly known EST without extending the known EST by at least 40 nucleotides in the 5′ direction were designated "EST." Of the sequences in the SignalTag™ database, 599 of the 5′ ESTs having a Von Heijne's score of at least 3.5 fell into this category.

Those 5′ ESTs matching a human mRNA sequence but extending the known sequence by at least 40 nucleotides in the 5′ direction were designated "VERT-ext." Of the sequences in the SignalTag™ database, 23 of the 5′ ESTs having a Von Heijne's score of at

least 3.5 fell into this category. Included in this category was a 5′ EST which extended the known sequence of the human translocase mRNA by more than 200 bases in the 5′ direction. A 5′ EST which extended the sequence of a human tumor suppressor gene in the 5′ direction was also identified.

Table V shows the distribution of 5′ ESTs in each category and the number of 5′ ESTs in each category having a given minimum von Heijne's score.

3. Evaluation of Spatial and Temporal Expression of mRNAs Corresponding to the 5'ESTs or Extended cDNAs

Each of the 5′ ESTs was also categorized based on the tissue from which its corresponding mRNA was obtained, as described below in Example 25.

EXAMPLE 25

Categorization of Expression Patterns

Table VI shows the distribution of 5' ESTs in each of the above defined category with respect to the tissue from which the 5'ESTs of the corresponding mRNA were obtained.

Table II provides the sequence identification numbers of 5′ EST sequences derived from different tissues, the categories in which these sequences fall, and the von Heijne's score of the signal peptides which they encode. The 5′ EST sequences and the amino acid sequences they encode are provided in the appended sequence listings. Table III provides the sequence ID numbers of the 5′ ESTs and the sequences of the signal peptides which they encode. The sequences of the 5′ ESTs and the polypeptides they encode are provided in the sequence listing appended hereto.

The sequences of DNA SEQ ID NOs: 38-291 can readily be screened for any errors therein and any sequence ambiguities can be resolved by resequencing a fragment containing such errors or amibiguities on both strands. Such fragments may be obtained from the plasmids stored in the inventors' laboratory or can be isolated using the techniques described herein. Resolution of any such ambiguities or errors may be facilitated by using primers which hybridize to sequences located close to the ambiguous or erroneous sequences. For example, the primers may hybridize to sequences within 50-75 bases of the amibiguity or

error. Upon resolution of an error or ambiguity, the corresponding corrections can be made in the protein sequences encoded by the DNA containing the error or amibiguity.

In addition to categorizing the 5′ ESTs with respect to their tissue of origin, the spatial and temporal expression patterns of the mRNAs corresponding to the 5′ ESTs, as well as their expression levels, may be determined as described in Example 26 below. Characterization of the spatial and temporal expression patterns and expression levels of these mRNAs is useful for constructing expression vectors capable of producing a desired level of gene product in a desired spatial or temporal manner, as will be discussed in more detail below.

Furthermore, 5′ ESTs whose corresponding mRNAs are associated with disease states may also be identified. For example, a particular disease may result from the lack of expression, over expression, or under expression of an mRNA corresponding to a 5′ EST. By comparing mRNA expression patterns and quantities in samples taken from healthy individuals with those from individuals suffering from a particular disease, 5′ ESTs responsible for the disease may be identified.

It will be appreciated that the results of the above characterization procedures for 5′ ESTs also apply to extended cDNAs (obtainable as described below) which contain sequences adjacent to the 5′ ESTs. It will also be appreciated that if desired, characterization may be delayed until extended cDNAs have been obtained rather than characterizing the ESTs themselves.

EXAMPLE 26

Evaluation of Expression Levels and Patterns of mRNAs Corresponding to 5′ ESTs or Extended cDNAs

Expression levels and patterns of mRNAs corresponding to 5′ ESTs or extended cDNAs (obtainable as described below in example 27) may be analyzed by solution hybridization with long probes as described in International Patent Application No. WO 97/05277, the entire contents of which are hereby incorporated by reference. Briefly, a 5′ EST, extended cDNA, or fragment thereof corresponding to the gene encoding the mRNA to be characterized is inserted at a cloning site immediately downstream of a bacteriophage (T3,

T7 or SP6) RNA polymerase promoter to produce antisense RNA. Preferably, the 5′ EST or extended cDNA has 100 or more nucleotides. The plasmid is linearized and transcribed in the presence of ribonucleotides comprising modified ribonucleotides (*i.e.* biotin-UTP and DIG-UTP). An excess of this doubly labeled RNA is hybridized in solution with mRNA isolated from cells or tissues of interest. The hybridizations are performed under standard stringent conditions (40-50°C for 16 hours in an 80% formamide, 0.4 M NaCl buffer, pH 7-8). The unhybridized probe is removed by digestion with ribonucleases specific for single-stranded RNA (*i.e.* RNases CL3, T1, Phy M, U2 or A). The presence of the biotin-UTP modification enables capture of the hybrid on a microtitration plate coated with streptavidin. The presence of the DIG modification enables the hybrid to be detected and quantified by ELISA using an anti-DIG antibody coupled to alkaline phosphatase.

The 5′ ESTs, extended cDNAs, or fragments thereof may also be tagged with nucleotide sequences for the serial analysis of gene expression (SAGE) as disclosed in UK Patent Application No. 2 305 241 A, the entire contents of which are incorporated by reference. In this method, cDNAs are prepared from a cell, tissue, organism or other source of nucleic acid for which gene expression patterns must be determined. The resulting cDNAs are separated into two pools. The cDNAs in each pool are cleaved with a first restriction endonuclease, called an anchoring enzyme, having a recognition site which is likely to be present at least once in most cDNAs. The fragments which contain the 5′ or 3′ most region of the cleaved cDNA are isolated by binding to a capture medium such as streptavidin coated beads. A first oligonucleotide linker having a first sequence for hybridization of an amplification primer and an internal restriction site for a so-called tagging endonuclease is ligated to the digested cDNAs in the first pool. Digestion with the second endonuclease produces short tag fragments from the cDNAs.

A second oligonucleotide having a second sequence for hybridization of an amplification primer and an internal restriction site is ligated to the digested cDNAs in the second pool. The cDNA fragments in the second pool are also digested with the tagging endonuclease to generate short tag fragments derived from the cDNAs in the second pool. The tags resulting from digestion of the first and second pools with the anchoring enzyme and the tagging endonuclease are ligated to one another to produce so-called ditags. In some embodiments, the ditags are concatamerized to produce ligation products containing from 2

to 200 ditags. The tag sequences are then determined and compared to the sequences of the 5′ ESTs or extended cDNAs to determine which 5′ ESTs or extended cDNAs are expressed in the cell, tissue, organism, or other source of nucleic acids from which the tags were derived. In this way, the expression pattern of the 5′ ESTs or extended cDNAs in the cell, tissue, organism, or other source of nucleic acids is obtained.

Quantitative analysis of gene expression may also be performed using arrays. As used herein, the term array means a one dimensional, two dimensional, or multidimensional arrangement of full length cDNAs (*i.e.* extended cDNAs which include the coding sequence for the signal peptide, the coding sequence for the mature protein, and a stop codon), extended cDNAs, 5′ ESTs or fragments thereof of sufficient length to permit specific detection of gene expression. Preferably, the fragments are at least 15 nucleotides in length. More preferably, the fragments are at least 100 nucleotide long. More preferably, the fragments are more than 100 nucleotides in length. In some embodiments, the fragments may be more than 500 nucleotide long.

For example, quantitative analysis of gene expression may be performed with full length cDNAs as defined below, extended cDNAs, 5′ ESTs, or fragments thereof in a complementary DNA microarray as described by Schena *et al.* (*Science* 270:467-470, 1995; *Proc. Natl. Acad. Sci. U.S.A.* 93:10614-10619, 1996). Full length cDNAs, extended cDNAs, 5′ ESTs or fragments thereof are amplified by PCR and arrayed from 96-well microtiter plates onto silylated microscope slides using high-speed robotics. Printed arrays are incubated in a humid chamber to allow rehydration of the array elements and rinsed, once in 0.2% SDS for 1 min, twice in water for 1 min and once for 5 min in sodium borohydride solution. The arrays are submerged in water for 2 min at 95°C, transferred into 0.2% SDS for 1 min, rinsed twice with water, air dried and stored in the dark at 25°C.

Cell or tissue mRNA is isolated or commercially obtained and probes are prepared by a single round of reverse transcription. Probes are hybridized to 1 cm^2 microarrays under a 14 x 14 mm glass coverslip for 6-12 hours at 60°C. Arrays are washed for 5 min at 25°C in low stringency wash buffer (1 x SSC/0.2% SDS), then for 10 min at room temperature in high stringency wash buffer (0.1 x SSC/0.2% SDS). Arrays are scanned in 0.1 x SSC using a fluorescence laser scanning device fitted with a custom filter set. Accurate differential

expression measurements are obtained by taking the average of the ratios of two independent hybridizations.

Quantitative analysis of the expression of genes may also be performed with full length cDNAs, extended cDNAs, 5′ ESTs, or fragments thereof in complementary DNA arrays as described by Pietu *et al.*. (*Genome Research* 6:492-503, 1996). The full length cDNAs, extended cDNAs, 5′ ESTs or fragments thereof are PCR amplified and spotted on membranes. Then, mRNAs originating from various tissues or cells are labeled with radioactive nucleotides. After hybridization and washing in controlled conditions, the hybridized mRNAs are detected by phospho-imaging or autoradiography. Duplicate experiments are performed and a quantitative analysis of differentially expressed mRNAs is then performed.

Alternatively, expression analysis of the 5′ ESTs or extended cDNAs can be done through high density nucleotide arrays as described by Lockhart *et al.* (*Nature Biotechnology* 14: 1675-1680, 1996) and Sosnowsky *et al.* (*Proc. Natl. Acad. Sci.* 94:1119-1123, 1997). Oligonucleotides of 15-50 nucleotides corresponding to sequences of the 5′ ESTs or extended cDNAs are synthesized directly on the chip (Lockhart *et al.*, *supra*) or synthesized and then addressed to the chip (Sosnowsky *et al.*, *supra*). Preferably, the oligonucleotides are about 20 nucleotides in length.

cDNA probes labeled with an appropriate compound, such as biotin, digoxigenin or fluorescent dye, are synthesized from the appropriate mRNA population and then randomly fragmented to an average size of 50 to 100 nucleotides. The said probes are then hybridized to the chip. After washing as described in Lockhart *et al*, *supra* and application of different electric fields (Sonowsky et *al*, *supra*.), the dyes or labeling compounds are detected and quantified. Duplicate hybridizations are performed. Comparative analysis of the intensity of the signal originating from cDNA probes on the same target oligonucleotide in different cDNA samples indicates a differential expression of the mRNA corresponding to the 5′ EST or extended cDNA from which the oligonucleotide sequence has been designed.

III. Use of 5′ ESTs to Clone Extended cDNAs and to Clone the Corresponding Genomic DNAs

Once 5′ ESTs which include the 5′ end of the corresponding mRNAs have been selected using the procedures described above, they can be utilized to isolate extended cDNAs which contain sequences adjacent to the 5′ ESTs. The extended cDNAs may include the entire coding sequence of the protein encoded by the corresponding mRNA, including the authentic translation start site, the signal sequence, and the sequence encoding the mature protein remaining after cleavage of the signal peptide. Such extended cDNAs are referred to herein as "full length cDNAs." Alternatively, the extended cDNAs may include only the sequence encoding the mature protein remaining after cleavage of the signal peptide, or only the sequence encoding the signal peptide.

Example 27 below describes a general method for obtaining extended cDNAs using 5' ESTs. Example 28 below provides experimental results, using the method explained in example 27, describing several extended cDNAs including the entire coding sequence and authentic 5′ end of the corresponding mRNA for several secreted proteins.

The methods of Examples 27, 28, and 29 can also be used to obtain extended cDNAs which encode less than the entire coding sequence of the secreted proteins encoded by the genes corresponding to the 5′ ESTs. In some embodiments, the extended cDNAs isolated using these methods encode at least 10 amino acids of one of the proteins encoded by the sequences of SEQ ID NOs: 38-291. In further embodiments, the extended cDNAs encode at least 20 amino acids of the proteins encoded by the sequences of SEQ ID NOs: 38-291. In further embodiments, the extended cDNAs encode at least 30 amino amino acids of the sequences of SEQ ID NOs: 38-291. In a preferred embodiment, the extended cDNAs encode a full length protein sequence, which includes the protein coding sequences of SEQ ID NOs: 38-291.

EXAMPLE 27

General Method for Using 5′ ESTs to Clone and Sequence cDNAs which Include the Entire Coding Region and the Authentic 5′ End of the Corresponding mRNA

The following general method has been used to quickly and efficiently isolate extended cDNAs having the authentic 5′ ends of their corresponding mRNAs as well as

the full protein coding sequence and including sequence adjacent to the sequences of the 5′ ESTs used to obtain them. This method may be applied to obtain extended cDNAs for any 5′ EST in the NetGene™ database, including those 5′ ESTs encoding polypeptides belonging to secreted proteins. The method is summarized in figure 3.

1. Obtention of Extended cDNAs

a) First strand synthesis

The method takes advantage of the known 5′ sequence of the mRNA. A reverse transcription reaction is conducted on purified mRNA with a poly 14dT primer containing a 49 nucleotide sequence at its 5′ end allowing the addition of a known sequence at the end of the cDNA which corresponds to the 3′ end of the mRNA. For example, the primer may have the following sequence: 5′-ATC GTT GAG ACT CGT ACC AGC AGA GTC ACG AGA GAG ACT ACA CGG TAC TGG TTT TTT TTT TTT TTVN -3′ (SEQ ID NO:14). Those skilled in the art will appreciate that other sequences may also be added to the poly dT sequence and used to prime the first strand synthesis. Using this primer and a reverse transcriptase such as the Superscript II (Gibco BRL) or Rnase H Minus M-MLV (Promega) enzyme, a reverse transcript anchored at the 3′ polyA site of the RNAs is generated.

After removal of the mRNA hybridized to the first cDNA strand by alkaline hydrolysis, the products of the alkaline hydrolysis and the residual poly dT primer are eliminated with an exclusion column such as an AcA34 (Biosepra) matrix as explained in Example 11.

b) Second strand synthesis

A pair of nested primers on each end is designed based on the known 5′ sequence from the 5′ EST and the known 3′ end added by the poly dT primer used in the first strand synthesis. Softwares used to design primers are either based on GC content and melting temperatures of oligonucleotides, such as OSP (Illier and Green, *PCR Meth. Appl.* 1:124-128, 1991), or based on the octamer frequency disparity method (Griffais *et al.*, *Nucleic Acids Res.* 19: 3887-3891, 1991) such as PC-Rare (http://bioinformatics.weizmann.ac.il/software/PC-Rare/doc/manuel.html).

Preferably, the nested primers at the 5′ end are separated from one another by four to nine bases. The 5′ primer sequences may be selected to have melting temperatures and specificities suitable for use in PCR.

Preferably, the nested primers at the 3′ end are separated from one another by four to nine bases. For example, the nested 3′ primers may have the following sequences: (5′- CCA GCA GAG TCA CGA GAG AGA CTA CAC GG -3′(SEQ ID NO:15), and 5′- CAC GAG AGA GAC TAC ACG GTA CTG G -3′ (SEQ ID NO:16). These primers were selected because they have melting temperatures and specificities compatible with their use in PCR. However, those skilled in the art will appreciate that other sequences may also be used as primers.

The first PCR run of 25 cycles is performed using the Advantage Tth Polymerase Mix (Clontech) and the outer primer from each of the nested pairs. A second 20 cycle PCR using the same enzyme and the inner primer from each of the nested pairs is then performed on 1/2500 of the first PCR product. Thereafter, the primers and nucleotides are removed.

2. Sequencing of Full Length Extended cDNAs or Fragments Thereof

Due to the lack of position constraints on the design of 5' nested primers compatible for PCR use using the OSP software, amplicons of two types are obtained. Preferably, the second 5' primer is located upstream of the translation initiation codon thus yielding a nested PCR product containing the whole coding sequence. Such a full length extended cDNA undergoes a direct cloning procedure as described in section a. However, in some cases, the second 5' primer is located downstream of the translation initiation codon, thereby yielding a PCR product containing only part of the ORF. Such incomplete PCR products are submitted to a modified procedure described in section b.

a) Nested PCR products containing complete ORFs

When the resulting nested PCR product contains the complete coding sequence, as predicted from the 5'EST sequence, it is cloned in an appropriate vector such as pED6dpc2, as described in section 3.

b) Nested PCR products containing incomplete ORFs

When the amplicon does not contain the complete coding sequence, intermediate steps are necessary to obtain both the complete coding sequence and a PCR product

containing the full coding sequence. The complete coding sequence can be assembled from several partial sequences determined directly from different PCR products as described in the following section.

Once the full coding sequence has been completely determined, new primers compatible for PCR use are designed to obtain amplicons containing the whole coding region. However, in such cases, 3' primers compatible for PCR use are located inside the 3' UTR of the corresponding mRNA, thus yielding amplicons which lack part of this region, *i.e.* the polyA tract and sometimes the polyadenylation signal, as illustrated in figure 3. Such full length extended cDNAs are then cloned into an appropriate vector as described in section 3.

c) Sequencing extended cDNAs

Sequencing of extended cDNAs is performed using a Die Terminator approach with the AmpliTaq DNA polymerase FS kit available from Perkin Elmer.

In order to sequence PCR fragments, primer walking is performed using software such as OSP to choose primers and automated computer software such as ASMG (Sutton *et al.*, *Genome Science Technol.* 1: 9-19, 1995) to construct contigs of walking sequences including the initial 5' tag using minimum overlaps of 32 nucleotides. Preferably, primer walking is performed until the sequences of full length cDNAs are obtained.

Completion of the sequencing of a given extended cDNA fragment is assessed as follows. Since sequences located after a polyA tract are difficult to determine precisely in the case of uncloned products, sequencing and primer walking processes for PCR products are interrupted when a polyA tract is identified in extended cDNAs obtained as described in case b. The sequence length is compared to the size of the nested PCR product obtained as described above. Due to the limited accuracy of the determination of the PCR product size by gel electrophoresis, a sequence is considered complete if the size of the obtained sequence is at least 70 % the size of the first nested PCR product. If the length of the sequence determined from the computer analysis is not at least 70 % of the length of the nested PCR product, these PCR products are cloned and the sequence of the insertion is determined. When Northern blot data are available, the size of the mRNA detected for a given PCR product is used to finally assess that the sequence is complete. Sequences which do not fulfill the above criteria are discarded and will undergo a new isolation procedure.

Sequence data of all extended cDNAs are then transferred to a proprietary database, where quality controls and validation steps are carried out as described in example 15.

3. Cloning of Full Length Extended cDNAs

The PCR product containing the full coding sequence is then cloned in an appropriate vector. For example, the extended cDNAs can be cloned into the expression vector pED6dpc2 (DiscoverEase, Genetics Institute, Cambridge, MA) as follows. pED6dpc2 vector DNA is prepared with blunt ends by performing an EcoRI digestion followed by a fill in reaction. The blunt ended vector is dephosphorylated. After removal of PCR primers and ethanol precipitation, the PCR product containing the full coding sequence or the extended cDNA obtained as described above is phosphorylated with a kinase subsequently removed by phenol-Sevag extraction and precipitation. The double stranded extended cDNA is then ligated to the vector and the resulting expression plasmid introduced into appropriate host cells.

Since the PCR products obtained as described above are blunt ended molecules that can be cloned in either direction, the orientation of several clones for each PCR product is determined. Then, 4 to 10 clones are ordered in microtiter plates and subjected to a PCR reaction using a first primer located in the vector close to the cloning site and a second primer located in the portion of the extended cDNA corresponding to the 3′ end of the mRNA. This second primer may be the antisense primer used in anchored PCR in the case of direct cloning (case a) or the antisense primer located inside the 3'UTR in the case of indirect cloning (case b). Clones in which the start codon of the extended cDNA is operably linked to the promoter in the vector so as to permit expression of the protein encoded by the extended cDNA are conserved and sequenced. In addition to the ends of cDNA inserts, approximately 50 bp of vector DNA on each side of the cDNA insert are also sequenced.

The cloned PCR products are then entirely sequenced according to the aforementioned procedure. In this case, contigation of long fragments is then performed on walking sequences that have already contigated for uncloned PCR products during primer walking. Sequencing of cloned amplicons is complete when the resulting contigs include the whole coding region as well as overlapping sequences with vector DNA on both ends.

4. Computer analysis of Full Length Extended cDNA

Sequences of all full length extended cDNAs are then submitted to further analysis as described below. Before searching the extended full length cDNAs for sequences of interest, extended cDNAs which are not of interest (vector RNAs, transfer RNAs, ribosomal RNAs, mitochondrial RNAs, prokaryotic RNAs and fungal RNAs) are discarded using methods essentially similar to those described for 5'ESTs in Example 18.

a) Identification of structural features

Structural features, e.g. polyA tail and polyadenylation signal, of the sequences of full length extended cDNAs are subsequently determined as follows.

A polyA tail is defined as a homopolymeric stretch of at least 11 A with at most one alternative base within it. The polyA tail search is restricted to the last 100 nt of the sequence and limited to stretches of 11 consecutive A's because sequencing reactions are often not readable after such a polyA stretch. Stretches having more than 90% homology over 8 nucleotides are identified as polyA tails using BLAST2N.

To search for a polyadenylation signal, the polyA tail is clipped from the full-length sequence. The 50 bp preceding the polyA tail are first searched for the canonic polyadenylation AAUAAA signal and, if the canonic signal is not detected, for the alternative AUUAAA signal (Sheets *et al., Nuc. Acids Res.* **18**: 5799-5805, 1990). If neither of these consensus polyadenylation signals is found, the canonic motif is searched again allowing one mismatch to account for possible sequencing errors. More than 85 % of identified polyadenylation signals of either type actually ends 10 to 30 bp from the polyA tail. Alternative AUUAAA signals represents approximately 15 % of the total number of identified polyadenylation signals.

b) Identification of functional features

Functional features, e.g. ORFs and signal sequences, of the sequences of full length extended cDNAs were subsequently determined as follows.

The 3 upper strand frames of extended cDNAs are searched for ORFs defined as the maximum length fragments beginning with a translation intiation codon and ending with a stop codon. ORFs encoding at least 20 amino acids are preferred.

Each found ORF is then scanned for the presence of a signal peptide in the first 50 amino-acids or, where appropriate, within shorter regions down to 20 amino acids or

less in the ORF, using the matrix method of von Heijne (*Nuc. Acids Res.* 14: 4683-4690, 1986), the disclosure of which is incorporated herein by reference as described in Example 22.

c) Homology to either nucleotidic or proteic sequences

Categorization of full-length sequences may be achieved using procedures essentially similar to those described for 5'ESTs in Example 24.

Extended cDNAs prepared as described above may be subsequently engineered to obtain nucleic acids which include desired portions of the extended cDNA using conventional techniques such as subcloning, PCR, or *in vitro* oligonucleotide synthesis. For example, nucleic acids which include only the full coding sequences (*i.e.* the sequences encoding the signal peptide and the mature protein remaining after the signal peptide is cleaved off) may be obtained using techniques known to those skilled in the art. Alternatively, conventional techniques may be applied to obtain nucleic acids which contain only the coding sequences for the mature protein remaining after the signal peptide is cleaved off or nucleic acids which contain only the coding sequences for the signal peptides.

Similarly, nucleic acids containing any other desired portion of the coding sequences for the secreted protein may be obtained. For example, the nucleic acid may contain at least 10 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. In another embodiment, the nucleic acid may contain at least 15 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. Alternatively, the nucleic acid may contain at least 20 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. In another embodiment, the nucleic acid may contain at least 25 consecutive bases of an extended cDNA such as one of the extended cDNAs described below. In yet another embodiment, the nucleic acid may contain at least 40 consecutive bases of an extended cDNA such as one of the extended cDNAs described below.

Once an extended cDNA has been obtained, it can be sequenced to determine the amino acid sequence it encodes. Once the encoded amino acid sequence has been determined, one can create and identify any of the many conceivable cDNAs that will encode that protein by simply using the degeneracy of the genetic code. For example, allelic variants

or other homologous nucleic acids can be identified as described below. Alternatively, nucleic acids encoding the desired amino acid sequence can be synthesized *in vitro*.

In a preferred embodiment, the coding sequence may be selected using the known codon or codon pair preferences for the host organism in which the cDNA is to be expressed.

The extended cDNAs derived from the 5' ESTS of the present invention were obtained as described in Example 28 below.

EXAMPLE 28

Characterization of cloned extended cDNAs obtained using 5' ESTs

The procedure described in Example 27 above was used to obtain the extended cDNAs derived from the 5' ESTs of the present invention in a variety of tissues. The following list provides a few examples of thus obtained extended cDNAs.

Using this approach, the full length cDNA of SEQ ID NO:17 (internal identification number 48-19-3-G1-FL1) was obtained. This cDNA falls into the "EST-ext" category described above and encodes the signal peptide MKKVLLLITAILAVAVG (SEQ ID NO: 18) having a von Heijne score of 8.2.

The full length cDNA of SEQ ID NO:19 (internal identification number 58-34-2-E7-FL2) was also obtained using this procedure. This cDNA falls into the "EST-ext" category described above and encodes the signal peptide MWWFQQGLSFLPSALVIWTSA (SEQ ID NO:20) having a von Heijne score of 5.5.

Another full length cDNA obtained using the procedure described above has the sequence of SEQ ID NO:21 (internal identification number 51-27-1-E8-FL1). This cDNA, falls into the "EST-ext" category described above and encodes the signal peptide MVLTTLPSANSANSPVNMPTTGPNSLSYASSALSPCLT (SEQ ID NO:22) having a von Heijne score of 5.9.

The above procedure was also used to obtain a full length cDNA having the sequence of SEQ ID NO:23 (internal identification number 76-4-1-G5-FL1). This cDNA falls into the "EST-ext" category described above and encodes the signal peptide ILSTVTALTFAXA (SEQ ID NO:24) having a von Heijne score of 5.5.

The full length cDNA of SEQ ID NO:25 (internal identification number 51-3-3-B10-FL3) was also obtained using this procedure. This cDNA falls into the "new" category

described above and encodes a signal peptide LVLTLCTLPLAVA (SEQ ID NO:26) having a von Heijne score of 10.1.

The full length cDNA of SEQ ID NO:27 (internal identification number 58-35-2-F10-FL2) was also obtained using this procedure. This cDNA falls into the "new" category described above and encodes a signal peptide LWLLFFLVTAIHA (SEQ ID NO:28) having a von Heijne score of 10.7.

Bacterial clones containing plasmids containing the full length cDNAs described above are presently stored in the inventor's laboratories under the internal identification numbers provided above. The inserts may be recovered from the stored materials by growing an aliquot of the appropriate bacterial clone in the appropriate medium. The plasmid DNA can then be isolated using plasmid isolation procedures familiar to those skilled in the art such as alkaline lysis minipreps or large scale alkaline lysis plasmid isolation procedures. If desired the plasmid DNA may be further enriched by centrifugation on a cesium chloride gradient, size exclusion chromatography, or anion exchange chromatography. The plasmid DNA obtained using these procedures may then be manipulated using standard cloning techniques familiar to those skilled in the art. Alternatively, a PCR can be done with primers designed at both ends of the cDNA insertion. The PCR product which corresponds to the cDNA can then be manipulated using standard cloning techniques familiar to those skilled in the art.

The polypeptides encoded by the extended cDNAs may be screened for the presence of known structural or functional motifs or for the presence of signatures, small amino acid sequences which are well conserved amongst the members of a protein family. The conserved regions have been used to derive consensus patterns or matrices included in the PROSITE data bank, in particular in the file prosite.dat (Release 13.0 of November 1995, located at http://expasy.hcuge.ch/sprot/prosite.html. Prosite_convert and prosite_scan programs (http://ulrec3.unil.ch/ftpserveur/prosite_scan) may be used to find signatures on the extended cDNAs.

For each pattern obtained with the prosite_convert program from the prosite.dat file, the accuracy of the detection on a new protein sequence may be assessed by evaluating the frequency of irrelevant hits on the population of human secreted proteins included in the data bank SWISSPROT. The ratio between the number of hits on shuffled proteins (with a window size of 20 amino acids) and the number of hits on native (unshuffled) proteins may be

used as an index. Every pattern for which the ratio is greater than 20% (one hit on shuffled proteins for 5 hits on native proteins) may be skipped during the search with prosite_scan. The program used to shuffle protein sequences (db_shuffled) and the program used to determine the statistics for each pattern in the protein data banks (prosite_statistics) are available on the ftp site http://ulrec3.unil.ch/ftpserveur/prosite_scan.

In addition to PCR based methods for obtaining extended cDNAs, traditional hybridization based methods may also be employed. These methods may also be used to obtain the genomic DNAs which encode the mRNAs from which the 5′ ESTs were derived, mRNAs corresponding to the extended cDNAs, or nucleic acids which are homologous to extended cDNAs or 5′ ESTs. Example 29 below provides examples of such methods.

EXAMPLE 29

Methods for Obtaining cDNAs which include the Entire Coding Region and the Authentic 5′End of the Corresponding mRNA

A full length cDNA library can be made using the strategies described in Examples 13, 14, 15, and 16 above by replacing the random nonamer used in Example 14 with an oligo-dT primer. For instance, the oligonucleotide of SEQ ID NO:14 may be used.

Alternatively, a cDNA library or genomic DNA library may be obtained from a commercial source or made using techniques familiar to those skilled in the art. Such cDNA or genomic DNA librairies may be used to isolate extended cDNAs obtained from 5′ EST or nucleic acids homologous to extended cDNAs or 5′ EST as follows. The cDNA library or genomic DNA library is hybridized to a detectable probe comprising at least 10 consecutive nucleotides from the 5′ EST or extended cDNA using conventional techniques. Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST or extended cDNA. More preferably, the probe comprises at least 20 to 30 consecutive nucleotides from the 5′ EST or extended cDNA. In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST or extended cDNA.

Techniques for identifying cDNA clones in a cDNA library which hybridize to a given probe sequence are disclosed in Sambrook *et al., Molecular Cloning: A Laboratory Manual*

2d Ed., Cold Spring Harbor Laboratory Press, 1989, the disclosure of which is incorporated herein by reference. The same techniques may be used to isolate genomic DNAs.

Briefly, cDNA or genomic DNA clones which hybridize to the detectable probe are identified and isolated for further manipulation as follows. A probe comprising at least 10 consecutive nucleotides from the 5′ EST or extended cDNA is labeled with a detectable label such as a radioisotope or a fluorescent molecule. Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST or extended cDNA. More preferably, the probe comprises 20 to 30 consecutive nucleotides from the 5′ EST or extended cDNA. In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST or extended cDNA.

Techniques for labeling the probe are well known and include phosphorylation with polynucleotide kinase, nick translation, *in vitro* transcription, and non radioactive techniques. The cDNAs or genomic DNAs in the library are transferred to a nitrocellulose or nylon filter and denatured. After blocking of non specific sites, the filter is incubated with the labeled probe for an amount of time sufficient to allow binding of the probe to cDNAs or genomic DNAs containing a sequence capable of hybridizing thereto.

By varying the stringency of the hybridization conditions used to identify extended cDNAs or genomic DNAs which hybridize to the detectable probe, extended cDNAS having different levels of homology to the probe can be identified and isolated as described below.

1. Identification of Extended cDNA or Genomic cDNA Sequences Having a High Degree of Homology to the Labeled Probe

To identify extended cDNAs or genomic DNAs having a high degree of homology to the probe sequence, the melting temperature of the probe may be calculated using the following formulas:

For probes between 14 and 70 nucleotides in length the melting temperature (Tm) is calculated using the formula: $Tm=81.5+16.6(\log [Na+])+0.41(\text{fraction } G+C)-(600/N)$ where N is the length of the probe.

If the hybridization is carried out in a solution containing formamide, the melting temperature may be calculated using the equation $Tm=81.5+16.6(\log [Na+])+0.41(\text{fraction } G+C)-(0.63\% \text{ formamide})-(600/N)$ where N is the length of the probe.

Prehybridization may be carried out in 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 μg denatured fragmented salmon sperm DNA or 6X SSC, 5X Denhardt's reagent, 0.5% SDS, 100 μg denatured fragmented salmon sperm DNA, 50% formamide. The formulas for SSC and Denhardt's solutions are listed in Sambrook *et al.*, *supra*.

Hybridization is conducted by adding the detectable probe to the prehybridization solutions listed above. Where the probe comprises double stranded DNA, it is denatured before addition to the hybridization solution. The filter is contacted with the hybridization solution for a sufficient period of time to allow the probe to hybridize to extended cDNAs or genomic DNAs containing sequences complementary thereto or homologous thereto. For probes over 200 nucleotides in length, the hybridization may be carried out at 15-25°C below the Tm. For shorter probes, such as oligonucleotide probes, the hybridization may be conducted at 15-25°C below the Tm. Preferably, for hybridizations in 6X SSC, the hybridization is conducted at approximately 68°C. Preferably, for hybridizations in 50% formamide containing solutions, the hybridization is conducted at approximately 42°C.

All of the foregoing hybridizations would be considered to be under "stringent" conditions.

Following hybridization, the filter is washed in 2X SSC, 0.1% SDS at room temperature for 15 minutes. The filter is then washed with 0.1X SSC, 0.5% SDS at room temperature for 30 minutes to 1 hour. Thereafter, the solution is washed at the hybridization temperature in 0.1X SSC, 0.5% SDS. A final wash is conducted in 0.1X SSC at room temperature.

Extended cDNAs, nucleic acids homologous to extended cDNAs or 5' ESTs, or genomic DNAs which have hybridized to the probe are identified by autoradiography or other conventional techniques.

<u>2. Obtention of Extended cDNA or Genomic cDNA Sequences Having Lower Degrees of Homology to the Labeled Probe</u>

The above procedure may be modified to identify extended cDNAs, nucleic acids homologous to extended cDNAs, or genomic DNAs having decreasing levels of homology to the probe sequence. For example, to obtain extended cDNAs, nucleic acids homologous to extended cDNAs, or genomic DNAs of decreasing homology to the detectable probe, less stringent conditions may be used. For example, the hybridization temperature may be

decreased in increments of 5°C from 68°C to 42°C in a hybridization buffer having a sodium concentration of approximately 1M. Following hybridization, the filter may be washed with 2X SSC, 0.5% SDS at the temperature of hybridization. These conditions are considered to be "moderate" conditions above 50°C and "low" conditions below 50°C.

Alternatively, the hybridization may be carried out in buffers, such as 6X SSC, containing formamide at a temperature of 42°C. In this case, the concentration of formamide in the hybridization buffer may be reduced in 5% increments from 50% to 0% to identify clones having decreasing levels of homology to the probe. Following hybridization, the filter may be washed with 6X SSC, 0.5% SDS at 50°C. These conditions are considered to be "moderate" conditions above 25% formamide and "low" conditions below 25% formamide.

Extended cDNAs, nucleic acids homologous to extended cDNAs, or genomic DNAs which have hybridized to the probe are identified by autoradiography.

3. Determination of the Degree of Homology Between the Obtained Extended cDNAs and the Labeled Probe

If it is desired to obtain nucleic acids homologous to extended cDNAs, such as allelic variants thereof or nucleic acids encoding proteins related to the proteins encoded by the extended cDNAs, the level of homology between the hybridized nucleic acid and the extended cDNA or 5′ EST used as the probe may be further determined using BLAST2N; parameters may be adapted depending on the sequence length and degree of homology studied. To determine the level of homology between the hybridized nucleic acid and the extended cDNA or 5′EST from which the probe was derived, the nucleotide sequences of the hybridized nucleic acid and the extended cDNA or 5′EST from which the probe was derived are compared. For example, using the above methods, nucleic acids having at least 95% nucleic acid homology to the extended cDNA or 5′EST from which the probe was derived may be obtained and identified. Similarly, by using progressively less stringent hybridization conditions one can obtain and identify nucleic acids having at least 90%, at least 85%, at least 80% or at least 75% homology to the extended cDNA or 5′EST from which the probe was derived.

To determine whether a clone encodes a protein having a given amount of homology to the protein encoded by the extended cDNA or 5′ EST, the amino acid sequence encoded by the extended cDNA or 5′ EST is compared to the amino acid sequence encoded by the

hybridizing nucleic acid. Homology is determined to exist when an amino acid sequence in the extended cDNA or 5′ EST is closely related to an amino acid sequence in the hybridizing nucleic acid. A sequence is closely related when it is identical to that of the extended cDNA or 5′ EST or when it contains one or more amino acid substitutions therein in which amino acids having similar characteristics have been substituted for one another. Using the above methods and algorithms such as FASTA with parameters depending on the sequence length and degree of homology studied, one can obtain nucleic acids encoding proteins having at least 95%, at least 90%, at least 85%, at least 80% or at least 75% homology to the proteins encoded by the extended cDNA or 5′EST from which the probe was derived.

In addition to the above described methods, other protocols are available to obtain extended cDNAs using 5' ESTs as outlined in the following paragraphs.

Extended cDNAs may be prepared by obtaining mRNA from the tissue, cell, or organism of interest using mRNA preparation procedures utilizing polyA selection procedures or other techniques known to those skilled in the art. A first primer capable of hybridizing to the polyA tail of the mRNA is hybridized to the mRNA and a reverse transcription reaction is performed to generate a first cDNA strand.

The first cDNA strand is hybridized to a second primer containing at least 10 consecutive nucleotides of the sequences of SEQ ID NOs 38-291. Preferably, the primer comprises at least 12, 15, or 17 consecutive nucleotides from the sequences of SEQ ID NOs 38-291. More preferably, the primer comprises 20 to 30 consecutive nucleotides from the sequences of SEQ ID NOs 38-291. In some embodiments, the primer comprises more than 30 nucleotides from the sequences of SEQ ID NOs 38-291. If it is desired to obtain extended cDNAs containing the full protein coding sequence, including the authentic translation initiation site, the second primer used contains sequences located upstream of the translation initiation site. The second primer is extended to generate a second cDNA strand complementary to the first cDNA strand. Alternatively, RT-PCR may be performed as described above using primers from both ends of the cDNA to be obtained.

Extended cDNAs containing 5′ fragments of the mRNA may be prepared by hybridizing an mRNA comprising the sequence of the 5'EST for which an extended cDNA is desired with a primer comprising at least 10 consecutive nucleotides of the sequences

complementary to the 5'EST and reverse transcribing the hybridized primer to make a first cDNA strand from the mRNAs. Preferably, the primer comprises at least 12, 15, or 17 consecutive nucleotides from the 5'EST. More preferably, the primer comprises 20 to 30 consecutive nucleotides from the 5'EST.

Thereafter, a second cDNA strand complementary to the first cDNA strand is synthesized. The second cDNA strand may be made by hybridizing a primer complementary to sequences in the first cDNA strand to the first cDNA strand and extending the primer to generate the second cDNA strand.

The double stranded extended cDNAs made using the methods described above are isolated and cloned. The extended cDNAs may be cloned into vectors such as plasmids or viral vectors capable of replicating in an appropriate host cell. For example, the host cell may be a bacterial, mammalian, avian, or insect cell.

Techniques for isolating mRNA, reverse transcribing a primer hybridized to mRNA to generate a first cDNA strand, extending a primer to make a second cDNA strand complementary to the first cDNA strand, isolating the double stranded cDNA and cloning the double stranded cDNA are well known to those skilled in the art and are described *in Current Protocols in Molecular Biology*, John Wiley and Sons, Inc. 1997 and Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, 1989, the entire disclosures of which are incorporated herein by reference.

Alternatively, procedures such as the one described in Example 29 may be used for obtaining full length cDNAs or extended cDNAs. In this approach, full length or extended cDNAs are prepared from mRNA and cloned into double stranded phagemids as follows. The cDNA library in the double stranded phagemids is then rendered single stranded by treatment with an endonuclease, such as the Gene II product of the phage F1, and an exonuclease (Chang *et al.*, *Gene* **127**:95-8, 1993). A biotinylated oligonucleotide comprising the sequence of a 5' EST, or a fragment containing at least 10 nucleotides thereof, is hybridized to the single stranded phagemids. Preferably, the fragment comprises at least 12, 15, or 17 consecutive nucleotides from the 5' EST. More preferably, the fragment comprises 20-30 consecutive nucleotides from the 5' EST. In some procedures, the fragment may comprise more than 30 consecutive nucleotides from the 5' EST.

Hybrids between the biotinylated oligonucleotide and phagemids having inserts containing the 5′ EST sequence are isolated by incubating the hybrids with streptavidin coated paramagnetic beads and retrieving the beads with a magnet (Fry *et al.*, *Biotechniques*, **13**: 124-131, 1992). Therafter, the resulting phagemids containing the 5′ EST sequence are released from the beads and converted into double stranded DNA using a primer specific for the 5′ EST sequence. Alternatively, protocoles such as the Gene Trapper kit (Gibco BRL) may be used. The resulting double stranded DNA is transformed into bacteria. Extended cDNAs containing the 5′ EST sequence are identified by colony PCR or colony hybridization.

Using any of the above described methods in section III, a plurality of extended cDNAs containing full length protein coding sequences or sequences encoding only the mature protein remaining after the signal peptide is cleaved off may be provided as cDNA libraries for subsequent evaluation of the encoded proteins or use in diagnostic assays as described below.

IV. Expression of Proteins Encoded by Extended cDNAs Isolated Using 5′ ESTs

Extended cDNAs containing the full protein coding sequences of their corresponding mRNAs or portions thereof, such as cDNAs encoding the mature protein, may be used to express the encoded secreted proteins or portions thereof as described in Example 30 below. If desired, the extended cDNAs may contain the sequences encoding the signal peptide to facilitate secretion of the expressed protein. It will be appreciated that a plurality of extended cDNAs containing the full protein coding sequences or portions thereof may be simultaneously cloned into expression vectors to create an expression library for analysis of the encoded proteins as described below.

EXAMPLE 30

Expression of the Proteins Encoded by the Genes Corresponding to 5'ESTS or Portions Thereof

To express the proteins encoded by the genes corresponding to 5′ ESTs (or portions thereof), full length cDNAs containing the entire protein coding region or extended cDNAs containing sequences adjacent to the 5′ ESTs (or portions thereof) are obtained as described

in Examples 27-29 and cloned into a suitable expression vector. If desired, the nucleic acids may contain the sequences encoding the signal peptide to facilitate secretion of the expressed protein. The nucleic acids inserted into the expression vectors may also contain sequences upstream of the sequences encoding the signal peptide, such as sequences which regulate expression levels or sequences which confer tissue specific expression.

The nucleic acid encoding the protein or polypeptide to be expressed is operably linked to a promoter in an expression vector using conventional cloning technology. The expression vector may be any of the mammalian, yeast, insect or bacterial expression systems known in the art. Commercially available vectors and expression systems are available from a variety of suppliers including Genetics Institute (Cambridge, MA), Stratagene (La Jolla, California), Promega (Madison, Wisconsin), and Invitrogen (San Diego, California). If desired, to enhance expression and facilitate proper protein folding, the codon context and codon pairing of the sequence may be optimized for the particular expression organism in which the expression vector is introduced, as explained by Hatfield, *et al.*, U.S. Patent No. 5,082,767, incorporated herein by this reference.

The cDNA cloned into the expression vector may encode the entire protein (*i.e.* the signal peptide and the mature protein), the mature protein (*i.e.* the protein created by cleaving the signal peptide off), only the signal peptide or any other portion thereof.

The following is provided as one exemplary method to express the proteins encoded by the extended cDNAs corresponding to the 5′ ESTs or the nucleic acids described above. First, the methionine initiation codon for the gene and the polyA signal of the gene are identified. If the nucleic acid encoding the polypeptide to be expressed lacks a methionine to serve as the initiation site, an initiating methionine can be introduced next to the first codon of the nucleic acid using conventional techniques. Similarly, if the extended cDNA lacks a polyA signal, this sequence can be added to the construct by, for example, splicing out the polyA signal from pSG5 (Stratagene) using BglII and SalI restriction endonuclease enzymes and incorporating it into the mammalian expression vector pXT1 (Stratagene). pXT1 contains the LTRs and a portion of the *gag* gene from Moloney Murine Leukemia Virus. The position of the LTRs in the construct allow efficient stable transfection. The vector includes the Herpes Simplex thymidine kinase promoter and the selectable neomycin gene. The extended cDNA or portion thereof encoding the polypeptide to be expressed is obtained

by PCR from the bacterial vector using oligonucleotide primers complementary to the extended cDNA or portion thereof and containing restriction endonuclease sequences for Pst I incorporated into the 5′primer and BglII at the 5′ end of the corresponding cDNA 3′ primer, taking care to ensure that the extended cDNA is positioned with the poly A signal. The purified fragment obtained from the resulting PCR reaction is digested with PstI, blunt ended with an exonuclease, digested with Bgl II, purified and ligated to pXT1 containing a poly A signal and prepared for this ligation (blunt/BglII).

The ligated product is transfected into mouse NIH 3T3 cells using Lipofectin (Life Technologies, Inc., Grand Island, New York) under conditions outlined in the product specification. Positive transfectants are selected after growing the transfected cells in 600 μg/ml G418 (Sigma, St. Louis, Missouri). Preferably the expressed protein is released into the culture medium, thereby facilitating purification.

Alternatively, the extended cDNAs may be cloned into pED6dpc2 as described above. The resulting pED6dpc2 constructs may be transfected into a suitable host cell, such as COS 1 cells. Methotrexate resistant cells are selected and expanded. Preferably, the protein expressed from the extended cDNA is released into the culture medium thereby facilitating purification.

Proteins in the culture medium are separated by gel electrophoresis. If desired, the proteins may be ammonium sulfate precipitated or separated based on size or charge prior to electrophoresis.

As a control, the expression vector lacking a cDNA insert is introduced into host cells or organisms and the proteins in the medium are harvested. The secreted proteins present in the medium are detected using techniques familiar to those skilled in the art such as Coomassie blue or silver staining or using antibodies against the protein encoded by the extended cDNA .

Antibodies capable of specifically recognizing the protein of interest may be generated using synthetic 15-mer peptides having a sequence encoded by the appropriate 5′ EST, extended cDNA, or portion thereof. The synthetic peptides are injected into mice to generate antibody to the polypeptide encoded by the 5′ EST, extended cDNA, or portion thereof.

Secreted proteins from the host cells or organisms containing an expression vector which contains the extended cDNA derived from a 5′ EST or a portion thereof are compared

to those from the control cells or organism. The presence of a band in the medium from the cells containing the expression vector which is absent in the medium from the control cells indicates that the extended cDNA encodes a secreted protein. Generally, the band corresponding to the protein encoded by the extended cDNA will have a mobility near that expected based on the number of amino acids in the open reading frame of the extended cDNA. However, the band may have a mobility different than that expected as a result of modifications such as glycosylation, ubiquitination, or enzymatic cleavage.

Alternatively, if the protein expressed from the above expression vectors does not contain sequences directing its secretion, the proteins expressed from host cells containing an expression vector with an insert encoding a secreted protein or portion thereof can be compared to the proteins expressed in control host cells containing the expression vector without an insert. The presence of a band in samples from cells containing the expression vector with an insert which is absent in samples from cells containing the expression vector without an insert indicates that the desired protein or portion thereof is being expressed. Generally, the band will have the mobility expected for the secreted protein or portion thereof. However, the band may have a mobility different than that expected as a result of modifications such as glycosylation, ubiquitination, or enzymatic cleavage.

The protein encoded by the extended cDNA may be purified using standard immunochromatography techniques. In such procedures, a solution containing the secreted protein, such as the culture medium or a cell extract, is applied to a column having antibodies against the secreted protein attached to the chromatography matrix. The secreted protein is allowed to bind the immunochromatography column. Thereafter, the column is washed to remove non-specifically bound proteins. The specifically bound secreted protein is then released from the column and recovered using standard techniques.

If antibody production is not possible, the extended cDNA sequence or portion thereof may be incorporated into expression vectors designed for use in purification schemes employing chimeric polypeptides. In such strategies, the coding sequence of the extended cDNA or portion thereof is inserted in frame with the gene encoding the other half of the chimera. The other half of the chimera may be β-globin or a nickel binding polypeptide. A chromatography matrix having antibody to β-globin or nickel attached thereto is then used to purify the chimeric protein. Protease cleavage sites may be engineered between the β-globin

gene or the nickel binding polypeptide and the extended cDNA or portion thereof. Thus, the two polypeptides of the chimera may be separated from one another by protease digestion.

One useful expression vector for generating β-globin chimerics is pSG5 (Stratagene), which encodes rabbit β-globin. Intron II of the rabbit β-globin gene facilitates splicing of the expressed transcript, and the polyadenylation signal incorporated into the construct increases the level of expression. These techniques as described are well known to those skilled in the art of molecular biology. Standard methods are published in methods texts such as Davis *et al.*, (*Basic Methods in Molecular Biology*, Davis, Dibner, and Battey, ed., Elsevier Press, NY, 1986) and many of the methods are available from Stratagene, Life Technologies, Inc., or Promega. Polypeptide may additionally be produced from the construct using *in vitro* translation systems such as the *In vitro* Express™ Translation Kit (Stratagene).

Following expression and purification of the secreted proteins encoded by the 5′ ESTs, extended cDNAs, or fragments thereof, the purified proteins may be tested for the ability to bind to the surface of various cell types as described in Example 31 below. It will be appreciated that a plurality of proteins expressed from these cDNAs may be included in a panel of proteins to be simultaneously evaluated for the activities specifically described below, as well as other biological roles for which assays for determining activity are available.

EXAMPLE 31

Analysis of Secreted Proteins to Determine Whether they Bind to the Cell Surface

The proteins encoded by the 5′ ESTs, extended cDNAs, or fragments thereof are cloned into expression vectors such as those described in Example 30. The proteins are purified by size, charge, immunochromatography or other techniques familiar to those skilled in the art. Following purification, the proteins are labeled using techniques known to those skilled in the art. The labeled proteins are incubated with cells or cell lines derived from a variety of organs or tissues to allow the proteins to bind to any receptor present on the cell surface. Following the incubation, the cells are washed to remove non-specifically bound protein. The labeled proteins are detected by autoradiography. Alternatively, unlabeled proteins may be incubated with the cells and detected with antibodies having a detectable label, such as a fluorescent molecule, attached thereto.

Specificity of cell surface binding may be analyzed by conducting a competition analysis in which various amounts of unlabeled protein are incubated along with the labeled protein. The amount of labeled protein bound to the cell surface decreases as the amount of competitive unlabeled protein increases. As a control, various amounts of an unlabeled protein unrelated to the labeled protein is included in some binding reactions. The amount of labeled protein bound to the cell surface does not decrease in binding reactions containing increasing amounts of unrelated unlabeled protein, indicating that the protein encoded by the cDNA binds specifically to the cell surface.

As discussed above, secreted proteins have been shown to have a number of important physiological effects and, consequently, represent a valuable therapeutic resource. The secreted proteins encoded by the extended cDNAs or portions thereof made according to Examples 27-29 may be evaluated to determine their physiological activities as described below.

EXAMPLE 32

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Cytokine, Cell Proliferation or Cell Differentiation Activity

As discussed above, secreted proteins may act as cytokines or may affect cellular proliferation or differentiation. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein encoded by the extended cDNAs is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M^+ (preB M^+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7c and CMK. The proteins encoded by the above extended cDNAs or portions thereof may be evaluated for their ability to regulate T cell or thymocyte proliferation in assays such as those described above or in the following references, which are incorporated herein by reference: *Current Protocols in Immunology*, Ed. by Coligan *et al.*., Greene Publishing Associates and Wiley-Interscience; Takai *et al. J. Immunol.* 137:3494-3500, 1986., Bertagnolli *et al.*, J. *Immunol.* 145:1706-1712, 1990., Bertagnolli *et al.*, *Cell.*

Immunol. 133:327-341, 1991; Bertagnolli, *et al., J. Immunol.* 149:3778-3783, 1992; Bowman *et al., J. Immunol.* 152:1756-1761, 1994.

In addition, numerous assays for cytokine production and/or the proliferation of spleen cells, lymph node cells and thymocytes are known. These include the techniques disclosed in *Current Protocols in Immunology, supra* 1:3.12.1-3.12.14; and Schreiber In *Current Protocols in Immunology, supra* 1 : 6.8.1-6.8.8.

The proteins encoded by the cDNAs may also be assayed for the ability to regulate the proliferation and differentiation of hematopoietic or lymphopoietic cells. Many assays for such activity are familiar to those skilled in the art, including the assays in the following references, which are incorporated herein by reference: Bottomly *et al.*, In *Current Protocols in Immunology., supra.* 1 : 6.3.1-6.3.12,; deVries *et al., J. Exp. Med.* 173:1205-1211, 1991; Moreau *et al., Nature* 36:690-692, 1988; Greenberger *et al., Proc. Natl. Acad. Sci. U.S.A.* 80:2931-2938, 1983; Nordan, R., In *Current Protocols in Immunology., supra.* 1 : 6.6.1-6.6.5; Smith et al., *Proc. Natl. Acad. Sci. U.S.A.* 83:1857-1861, 1986; Bennett *et al.*, in *Current Protocols in Immunology supra* 1 : 6.15.1; Ciarletta *et al.*, In *Current Protocols in Immunology. supra* 1 : 6.13.1.

The proteins encoded by the cDNAs may also be assayed for their ability to regulate T-cell responses to antigens. Many assays for such activity are familiar to those skilled in the art, including the assays described in the following references, which are incorporated herein by reference: Chapter 3 (*In Vitro* Assays for Mouse Lymphocyte Function), Chapter 6 (Cytokines and Their Cellular Receptors) and Chapter 7, (Immunologic Studies in Humans) in *Current Protocols in Immunology supra*; Weinberger *et al., Proc. Natl. Acad. Sci. USA* 77:6091-6095, 1980; Weinberger *et al., Eur. J. Immun.* 11:405-411, 1981; Takai *et al., J. Immunol.* 137:3494-3500, 1986; Takai *et al., J. Immunol.* 140:508-512, 1988.

Those proteins which exhibit cytokine, cell proliferation, or cell differentiation activity may then be formulated as pharmaceuticals and used to treat clinical conditions in which induction of cell proliferation or differentiation is beneficial. Alternatively, as described in more detail below, genes encoding these proteins or nucleic acids regulating the expression of these proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 33

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Activity as Immune System Regulators

The proteins encoded by the cDNAs may also be evaluated for their effects as immune regulators. For example, the proteins may be evaluated for their activity to influence thymocyte or splenocyte cytotoxicity. Numerous assays for such activity are familiar to those skilled in the art including the assays described in the following references, which are incorporated herein by reference: Chapter 3 (*In Vitro* Assays for Mouse Lymphocyte Function 3.1-3.19) and Chapter 7 (Immunologic studies in Humans) in *Current Protocols in Immunology* , Coligan *et al.*, Eds, Greene Publishing Associates and Wiley-Interscience; Herrmann *et al.*, *Proc. Natl. Acad. Sci. USA* 78:2488-2492, 1981; Herrmann *et al.*, *J. Immunol.* 128:1968-1974, 1982; Handa *et al.*, *J. Immunol.* 135:1564-1572, 1985; Takai *et al.*, *J. Immunol.* 137:3494-3500, 1986; Takai *et al.*, *J. Immunol.* 140:508-512, 1988; Bowman *et al.*, *J. Virology* 61:1992-1998; Bertagnolli *et al.*, *Cell. Immunol.* 133:327-341, 1991; Brown *et al.*, *J. Immunol.* 153:3079-3092, 1994.

The proteins encoded by the cDNAs may also be evaluated for their effects on T-cell dependent immunoglobulin responses and isotype switching. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Maliszewski, *J. Immunol.* 144:3028-3033, 1990; Mond *et al.* in *Current Protocols in Immunology*, 1 : 3.8.1-3.8.16, *supra*.

The proteins encoded by the cDNAs may also be evaluated for their effect on immune effector cells, including their effect on Th1 cells and cytotoxic lymphocytes. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Chapter 3 (*In Vitro* Assays for Mouse Lymphocyte Function 3.1-3.19) and Chapter 7 (Immunologic Studies in Humans) in *Current Protocols in Immunology, supra*; Takai *et al.*, *J. Immunol.* 137:3494-3500, 1986; Takai *et al.*, *J. Immunol.* 140:508-512, 1988; Bertagnolli *et al.*, *J. Immunol.* 149:3778-3783, 1992.

The proteins encoded by the cDNAs may also be evaluated for their effect on dendritic cell mediated activation of naive T-cells. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references,

which are incorporated herein by reference: Guery *et al.*, *J. Immunol.* 134:536-544, 1995; Inaba *et al.*, *J. Exp. Med.* 173:549-559, 1991; Macatonia *et al.*, *J. Immunol.* 154:5071-5079, 1995; Porgador *et al.J. Exp. Med* 182:255-260, 1995; Nair *et al.*, *J. Virol.* 67:4062-4069, 1993; Huang *et al.*, *Science* 264:961-965, 1994; Macatonia *et al.J. Exp. Med* 169:1255-1264, 1989; Bhardwaj *et al.*, *Journal of Clinical Investigation* 94:797-807, 1994; and Inaba *et al.*, *J. Exp. Med* 172:631-640, 1990.

The proteins encoded by the cDNAs may also be evaluated for their influence on the lifetime of lymphocytes. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Darzynkiewicz *et al.*, *Cytometry* 13:795-808, 1992; Gorczyca *et al.*, *Leukemia* 7:659-670, 1993; Gorczyca *et al.*, *Cancer Res.* 53:1945-1951, 1993; Itoh *et al.*, *Cell* 66:233-243, 1991; Zacharchuk, *J. Immunol.* 145:4037-4045, 1990; Zamai *et al.*, *Cytometry* 14:891-897, 1993; Gorczyca *et al.*, *Int. J. Oncol.* 1:639-648, 1992.

The proteins encoded by the cDNAs may also be evaluated for their influence on early steps of T-cell commitment and development. Numerous assays for such activity are familiar to those skilled in the art, including without limitation the assays disclosed in the following references, which are incorporated herein by references: Antica *et al.*, *Blood* 84:111-117, 1994; Fine *et al.*, *Cell. Immunol.* 155:111-122, 1994; Galy *et al.*, *Blood* 85:2770-2778, 1995; Toki *et al.*, *Proc. Nat. Acad Sci. USA* 88:7548-7551, 1991.

Those proteins which exhibit activity as immune system regulators activity may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of immune activity is beneficial. For example, the protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases caused by viral, bacterial, fungal or other infection may be treatable using a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., plamodium and various fungal infections such as candidiasis. Of course, in this regard, a protein encoded by

extended cDNAs derived from the 5' ESTs of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Alternatively, proteins encoded by extended cDNAs derived from the 5' ESTs of the present invention may be used in treatment of autoimmune disorders including, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention.

Using the proteins of the invention it may also be possible to regulate immune responses either up or down.

Down regulation may involve inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T-cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active non-antigen-specific process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after the end of exposure to the tolerizing agent. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions, such as, for example, B7 costimulation), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through

its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (e.g., B7-1, B7-3) or blocking antibody), prior to transplantation, can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, *Science* 257:789-792, 1992 and Turka *et al.*, *Proc. Natl. Acad. Sci USA*, 89:11102-11105, 1992. In addition, murine models of GVHD (see Paul ed., *Fundamental Immunology*, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor/ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which potentially involved in the disease process.

Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/pr/pr mice or NZB hybrid mice, murine autoimmuno collagen arthritis, diabetes mellitus in OD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., *supra*, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may involve either enhancing an existing immune response or eliciting an initial immune response as shown by the following examples. For instance, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory form of B lymphocyte antigens systemically.

Alternatively, antiviral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention or together with a stimulatory form of a soluble peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention and reintroducing the *in vitro* primed T cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to T cells *in vivo*, thereby activating the T cells.

In another application, upregulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (e.g., sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The

transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide encoded by extended cDNAs derived from the 5' ESTs of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules can be transfected with nucleic acids encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I α chain and β_2 microglobulin or an MHC class II α chain and an MHC class II β chain to thereby express MHC class I or MHC class II proteins on the cell surface, respectively. Expression of the appropriate MHC class I or class II molecules in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject. Alternatively, as described in more detail below, genes encoding these immune system regulator proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 34

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Hematopoiesis Regulating Activity

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their hematopoiesis regulating activity. For example, the effect of the proteins on embryonic stem cell differentiation may be evaluated. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following

references, which are incorporated herein by reference: Johansson *et al. Cell. Biol.* 15:141-151, 1995; Keller *et al.*, *Mol. Cell. Biol.* 13:473-486, 1993; McClanahan *et al.*, *Blood* 81:2903-2915, 1993.

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their influence on the lifetime of stem cells and stem cell differentiation. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Freshney, Methylcellulose Colony Forming Assays, in *Culture of Hematopoietic Cells.*, Freshney, *et al.*. Eds. pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama *et al.*, *Proc. Natl. Acad. Sci. USA* 89:5907-5911, 1992; McNiece and Briddell, in *Culture of Hematopoietic Cells, supra*; Neben *et al.*, *Exp. Hematol.* 22:353-359, 1994; Ploemacher and Cobblestone In *Culture of Hematopoietic Cells, supra*1-21, Spooncer *et al*, in *Culture of Hematopoietic Cells, supra*163-179 and Sutherland in *Culture of Hematopoietic Cells, supra.* 139-162.

Those proteins which exhibit hematopoiesis regulatory activity may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of hematopoeisis is beneficial, such as in the treatment of myeloid or lymphoid cell deficiencies. Involvement in regulating hematopoiesis is indicated even by marginal biological activity in support of colony forming cells or of factor-dependent cell lines. For example, proteins supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, indicates utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells. Proteins supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (*i.e.*, traditional CSF activity) may be useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression. Proteins supporting the growth and proliferation of megakaryocytes and consequently of platelets allows prevention or treatment of various platelet disorders such as thrombocytopenia, and generally may be used in place of or complementary to platelet transfusions. Proteins supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells may therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantion, including, without limitation, aplastic anemia and paroxysmal nocturnal

hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in vivo* or *ex vivo* (*i.e.*, in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy. Alternatively, as described in more detail below, genes encoding hematopoiesis regulating activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 35

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Regulation of Tissue Growth

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their effect on tissue growth. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in International Patent Publication No. WO95/16035, International Patent Publication No. WO95/05846 and International Patent Publication No. WO91/07491, which are incorporated herein by reference.

Assays for wound healing activity include, without limitation, those described in: Winter, *Epidermal Wound Healing*, pps. 71-112, Maibach and Rovee, eds., Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, *J. Invest. Dermatol.* 71:382-84, 1978, which are incorporated herein by reference.

Those proteins which are involved in the regulation of tissue growth may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of tissue growth is beneficial. For example, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the

improved fixation of artificial joints. *De novo* bone synthesis induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of bone-forming cell progenitors. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein encoded by extended cDNAs derived from the 5' ESTs of the present invention is tendon/ligament formation. A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. *De novo* tendon/ligament-like tissue formation induced by a composition encoded by extended cDNAs derived from the 5' ESTs of the present invention contributes to the repair of tendon or ligaments defects of congenital, traumatic or other origin and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions encoded by extended cDNAs derived from the 5' ESTs of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.*, for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium) muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to generate. A protein of the invention may also exhibit angiogenic activity.

A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokinc damage.

A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Alternatively, as described in more detail below, genes encoding tissue growth regulating activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 36

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Regulation of Reproductive Hormones

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their ability to regulate reproductive hormones, such as follicle stimulating hormone. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Vale *et al.*, *Endocrinol.* **91**:562-572, 1972; Ling *et al.*, *Nature* **321**:779-782, 1986; Vale *et al.*, *Nature* **321**:776-779, 1986; Mason *et al.*, *Nature* **318**:659-663, 1985; Forage *et al.*, *Proc. Natl. Acad. Sci. USA* **83**:3091-3095, 1986, Chapter 6.12 in *Current Protocols in Immunology*, Coligan *et al.* Eds. Greene Publishing Associates and Wiley-Intersciece ; Taub *et al.*, *J. Clin. Invest.* **95**:1370-1376, 1995; Lind *et al.*, *APMIS* **103**:140-146, 1995; Muller *et al.*, *Eur. J. Immunol.* **25**:1744-1748; Gruber *et al.*, *J. Immunol.* **152**:5860-5867, 1994; Johnston *et al.*, *J Immunol.* **153**:1762-1768, 1994.

Those proteins which exhibit activity as reproductive hormones or regulators of cell movement may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of reproductive hormones are beneficial. For example, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins are characterized by their ability to stimulate the release of FSH. Thus, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of

the inhibin-B group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885, the disclosure of which is incorporated herein by reference. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

Alternatively, as described in more detail below, genes encoding reproductive hormone regulating activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 37

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Chemotactic/Chemokinetic Activity

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for chemotactic/chemokinetic activity. For example, a protein encoded by extended cDNAs derived from the 5' ESTs of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: *Current Protocols in Immunology*, Ed by Coligan, Kruisbeek, Margulies, Shevach and Strober, Pub. Greene Publishing Associates and Wiley-Interscience, Chapter **6.12**: 6.12.1-6.12.28; Taub *et al., J. Clin. Invest.* **95**:1370-1376, 1995; Lind *et al., APMIS* **103**:140-146, 1995; Mueller *et al., Eur. J. Immunol.* **25**:1744-1748; Gruber *et al., J. Immunol.* **152**:5860-5867, 1994; Johnston *et al. J. Immunol.*, **153**:1762-1768, 1994.

EXAMPLE 38

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Regulation of Blood Clotting

The proteins encoded by the extended cDNAs or portions thereof may also be evaluated for their effects on blood clotting. Numerous assays for such activity are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Linet *et al., J. Clin. Pharmacol.* **26**:131-140, 1986; Burdick *et al., Thrombosis Res.* **45**:413-419, 1987; Humphrey *et al., Fibrinolysis* **5**:71-79, 1991; Schaub, *Prostaglandins* **35**:467-474, 1988.

Those proteins which are involved in the regulation of blood clotting may then be formulated as pharmaceuticals and used to treat clinical conditions in which regulation of blood clotting is beneficial. For example, a protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulations disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as infarction of cardiac and central nervous system

vessels (e.g., stroke)). Alternatively, as described in more detail below, genes encoding blood clotting activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 39

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Involvement in Receptor/Ligand Interactions

The proteins encoded by the extended cDNAs or a portion thereof may also be evaluated for their involvement in receptor/ligand interactions. Numerous assays for such involvement are familiar to those skilled in the art, including the assays disclosed in the following references, which are incorporated herein by reference: Chapter 7. 7.28.1-7.28.22 in *Current Protocols in Immunology*, Coligan *et al.* Eds. Greene Publishing Associates and Wiley-Interscience; Takai *et al.*, *Proc. Natl. Acad. Sci. USA* 84:6864-6868, 1987; Bierer *et al.*, *J. Exp. Med.* 168:1145-1156, 1988; Rosenstein *et al.*, *J. Exp. Med.* 169:149-160, 1989; Stoltenborg *et al.*, *J. Immunol. Methods* 175:59-68, 1994; Stitt *et al.*, *Cell* 80:661-670, 1995; Gyuris *et al.*, *Cell* 75:791-803, 1993.

For example, the proteins encoded by extended cDNAs derived from the 5' ESTs of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein encoded by extended cDNAs derived from the 5' ESTs of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions. Alternatively, as described in more detail below, genes encoding proteins involved in receptor/ligand

interactions or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 40

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Anti-Inflammatory Activity

The proteins encoded by the extended cDNAs or a portion thereof may also be evaluated for anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions, including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome), ischemia-reperfusioninury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine- or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Alternatively, as described in more detail below, genes encoding anti-inflammatory activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 41

Assaying the Proteins Expressed from Extended cDNAs or Portions Thereof for Tumor Inhibition Activity

The proteins encoded by the extended cDNAs or a portion thereof may also be evaluated for tumor inhibition activity. In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for

example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth. Alternatively, as described in more detail below, genes tumor inhibition activity proteins or nucleic acids regulating the expression of such proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein. Alternatively, as described in more detail below, genes encoding proteins involved in any of the above mentioned activities or nucleic acids regulating the expression of such

proteins may be introduced into appropriate host cells to increase or decrease the expression of the proteins as desired.

EXAMPLE 42

Identification of Proteins which Interact with Polypeptides Encoded by Extended cDNAs

Proteins which interact with the polypeptides encoded by cDNAs derived from the 5′ ESTs or fragments thereof, such as receptor proteins, may be identified using two hybrid systems such as the Matchmaker Two Hybrid System 2 (Catalog No. K1604-1, Clontech). As described in the manual accompanying the kit which is incorporated herein by reference, the the cDNAs derived from 5′ ESTs, or fragments thereof, are inserted into an expression vector such that they are in frame with DNA encoding the DNA binding domain of the yeast transcriptional activator GAL4. cDNAs in a cDNA library which encode proteins which might interact with the polypeptides encoded by the extended cDNAs or portions thereof are inserted into a second expression vector such that they are in frame with DNA encoding the activation domain of GAL4. The two expression plasmids are transformed into yeast and the yeast are plated on selection medium which selects for expression of selectable markers on each of the expression vectors as well as GAL4 dependent expression of the HIS3 gene. Transformants capable of growing on medium lacking histidine are screened for GAL4 dependent lacZ expression. Those cells which are positive in both the histidine selection and the lacZ assay contain plasmids encoding proteins which interact with the polypeptide encoded by the extended cDNAs or portions thereof.

Alternatively, the system described in Lustig *et al.*, *Methods in Enzymology* **283**: 83-99, 1997, and in U.S. Patent No. 5,654,150, the disclosure of which is incorporated herein by reference, may be used for identifying molecules which interact with the polypeptides encoded by extended cDNAs. In such systems, *in vitro* transcription reactions are performed on a pool of vectors containing extended cDNA inserts cloned downstream of a promoter which drives *in vitro* transcription. The resulting pools of mRNAs are introduced into *Xenopus laevis* oocytes. The oocytes are then assayed for a desired activity.

Alternatively, the pooled *in vitro* transcription products produced as described above may be translated *in vitro*. The pooled *in vitro* translation products can be assayed for a desired activity or for interaction with a known polypeptide.

Proteins or other molecules interacting with polypeptides encoded by extended cDNAs can be found by a variety of additional techniques. In one method, affinity columns containing the polypeptide encoded by the extended cDNA or a portion thereof can be constructed. In some versions, of this method the affinity column contains chimeric proteins in which the protein encoded by the extended cDNA or a portion thereof is fused to glutathione S-transferase. A mixture of cellular proteins or pool of expressed proteins as described above and is applied to the affinity column. Proteins interacting with the polypeptide attached to the column can then be isolated and analyzed on 2-D electrophoresis gel as described in Ramunsen *et al.*, *Electrophoresis* **18**:588-598, 1997, the disclosure of which is incorporated herein by reference. Alternatively, the proteins retained on the affinity column can be purified by electrophoresis based methods and sequenced. The same method can be used to isolate antibodies, to screen phage display products, or to screen phage display human antibodies.

Proteins interacting with polypeptides encoded by extended cDNAs or portions thereof can also be screened by using an Optical Biosensor as described in Edwards and Leatherbarrow, *Analytical Biochemistry* 246:1-6, 1997, the disclosure of which is incorporated herein by reference. The main advantage of the method is that it allows the determination of the association rate between the protein and other interacting molecules. Thus, it is possible to specifically select interacting molecules with a high or low association rate. Typically a target molecule is linked to the sensor surface (through a carboxymethl dextran matrix) and a sample of test molecules is placed in contact with the target molecules. The binding of a test molecule to the target molecule causes a change in the refractive index and/ or thickness. This change is detected by the Biosensor provided it occurs in the evanescent field (which extend a few hundred nanometers from the sensor surface). In these screening assays, the target molecule can be one of the polypeptides encoded by extended cDNAs or a portion thereof and the test sample can be a collection of proteins extracted from tissues or cells, a pool of expressed proteins, combinatorial peptide and/ or chemical libraries, or phage displayed peptides.

The tissues or cells from which the test proteins are extracted can originate from any species.

In other methods, a target protein is immobilized and the test population is a collection of unique polypeptides encoded by the extended cDNAs or portions thereof.

To study the interaction of the proteins encoded by the extended cDNAs or portions thereof with drugs, the microdialysis coupled to HPLC method described by Wang *et al.*, *Chromatographia* 44:205-208, 1997 or the affinity capillary electrophoresis method described by Busch *et al.*, *J. Chromatogr.* 777:311-328, 1997, the disclosures of which are incorporated herein by reference can be used.

It will be appreciated by those skilled in the art that the proteins expressed from the extended cDNAs or portions may be assayed for numerous activities in addition to those specifically enumerated above. For example, the expressed proteins may be evaluated for applications involving control and regulation of inflammation, tumor proliferation or metastasis, infection, or other clinical conditions. In addition, the proteins expressed from the extended cDNAs or portions thereof may be useful as nutritional agents or cosmetic agents.

The proteins expressed from the cDNAs or portions thereof may be used to generate antibodies capable of specifically binding to the expressed protein or fragments thereof as described in Example 40 below. The antibodies may capable of binding a full length protein encoded by a cDNA derived from a 5′ EST, a mature protein (*i.e.* the protein generated by cleavage of the signal peptide) encoded by a cDNA derived from a 5′ EST, or a signal peptide encoded by a cDNA derived from a 5′ EST. Alternatively, the antibodies may be capable of binding fragments of at least 10 amino acids of the proteins encoded by the above cDNAs. In some embodiments, the antibodies may be capable of binding fragments of at least 15 amino acids of the proteins encoded by the above cDNAs. In other embodiments, the antibodies may be capable of binding fragments of at least 25 amino acids of the proteins expressed from the extended cDNAs which comprise at least 25 amino acids of the proteins encoded by the above cDNAs. In further embodiments, the antibodies may be capable of binding fragments of at least 40 amino acids of the proteins encoded by the above cDNAs.

EXAMPLE 43

Production of an Antibody to a Human Protein

Substantially pure protein or polypeptide is isolated from the transfected or transformed cells as described in Example 30. The concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device, to the level of a few μg/ml. Monoclonal or polyclonal antibody to the protein can then be prepared as follows:

1. Monoclonal Antibody Production by Hybridoma Fusion

Monoclonal antibody to epitopes of any of the peptides identified and isolated as described can be prepared from murine hybridomas according to the classical method of Kohler, and Milstein, *Nature* 256:495, 1975 or derivative methods thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein or peptides derived therefrom over a period of a few weeks. The mouse is then sacrificed, and the antibody producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, *Meth. Enzymol.* 70:419, 1980, the disclosure of which is incorporated herein by reference and derivative methods thereof. Selected positive clones can be expanded and their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Davis *et al.* in *Basic Methods in Molecular Biology* Elsevier, New York. Section 21-2, the disclosure of which is incorporated herein by reference.

2. Polyclonal Antibody Production by Immunization

Polyclonal antiserum containing antibodies to heterogenous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein or peptides derived therefrom, which can be unmodified or modified to enhance

immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less immunogenic than others and may require the use of carriers and adjuvant. Also, host animals response vary depending on site of inoculations and doses, with both inadequate or excessive doses of antigen resulting in low titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appears to be most reliable. An effective immunization protocol for rabbits can be found in Vaitukaitis. *et al, J. Clin. Endocrinol. Metab.* 33:988-991 (1971) , the disclosure of which is incorporated herein by reference..

Booster injections can be given at regular intervals, and antiserum harvested when antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony, *et al.*, Chap. 19 in: *Handbook of Experimental Immunology* D. Wier (ed) Blackwell (1973) , the disclosure of which is incorporated herein by reference. Plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/ml of serum (about 12 μM). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, D., Chap. 42 in: *Manual of Clinical Immunology*, 2d Ed. (Rose and Friedman, Eds.) Amer. Soc. For Microbiol., Washington, D.C. (1980) , the disclosure of which is incorporated herein by reference..

Antibody preparations prepared according to either protocol are useful in quantitative immunoassays which determine concentrations of antigen-bearing substances in biological samples; they are also used semi-quantitatively or qualitatively to identify the presence of antigen in a biological sample. The antibodies may also be used in therapeutic compositions for killing cells expressing the protein or reducing the levels of the protein in the body.

V. Use of 5' ESTs or Sequences Obtainable Therefrom or Portions Thereof as Reagents

The 5' ESTs of the present invention (or cDNAs or genomic DNAs obtainable therefrom) may be used as reagents in isolation procedures, diagnostic assays, and forensic procedures. For example, sequences from the 5' ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be detectably labeled and used as probes to isolate

other sequences capable of hybridizing to them. In addition, sequences from 5' ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be used to design PCR primers to be used in isolation, diagnostic, or forensic procedures.

1. Use of 5' ESTs or Sequences Obtainable Therefrom or Portions Thereof in Isolation, Diagnostic and Forensic Procedures

EXAMPLE 44

Preparation of PCR Primers and Amplification of DNA

The 5' EST sequences (or cDNAs or genomic DNAs obtainable therefrom) may be used to prepare PCR primers for a variety of applications, including isolation procedures for cloning nucleic acids capable of hybridizing to such sequences, diagnostic techniques and forensic techniques. The PCR primers are at least 10 bases, and preferably at least 12, 15, or 17 bases in length. More preferably, the PCR primers are at least 20-30 bases in length. In some embodiments, the PCR primers may be more than 30 bases in length. It is preferred that the primer pairs have approximately the same G/C ratio, so that melting temperatures are approximately the same. A variety of PCR techniques are familiar to those skilled in the art. For a review of PCR technology, see Molecular Cloning to Genetic Engineering, White Ed. *in Methods in Molecular Biology* 67: Humana Press, Totowa 1997, the disclosure of which is incorporated herein by reference. In each of these PCR procedures, PCR primers on either side of the nucleic acid sequences to be amplified are added to a suitably prepared nucleic acid sample along with dNTPs and a thermostable polymerase such as Taq polymerase, Pfu polymerase, or Vent polymerase. The nucleic acid in the sample is denatured and the PCR primers are specifically hybridized to complementary nucleic acid sequences in the sample. The hybridized primers are extended. Thereafter, another cycle of denaturation, hybridization, and extension is initiated. The cycles are repeated multiple times to produce an amplified fragment containing the nucleic acid sequence between the primer sites.

EXAMPLE 45

Use of 5'ESTs as Probes

Probes derived from 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom), including full length cDNAs or genomic sequences, may be labeled with detectable labels familiar to those skilled in the art, including radioisotopes and non-radioactive labels, to provide a detectable probe. The detectable probe may be single stranded or double stranded and may be made using techniques known in the art, including *in vitro* transcription, nick translation, or kinase reactions. A nucleic acid sample containing a sequence capable of hybridizing to the labeled probe is contacted with the labeled probe. If the nucleic acid in the sample is double stranded, it may be denatured prior to contacting the probe. In some applications, the nucleic acid sample may be immobilized on a surface such as a nitrocellulose or nylon membrane. The nucleic acid sample may comprise nucleic acids obtained from a variety of sources, including genomic DNA, cDNA libraries, RNA, or tissue samples.

Procedures used to detect the presence of nucleic acids capable of hybridizing to the detectable probe include well known techniques such as Southern blotting, Northern blotting, dot blotting, colony hybridization, and plaque hybridization. In some applications, the nucleic acid capable of hybridizing to the labeled probe may be cloned into vectors such as expression vectors, sequencing vectors, or *in vitro* transcription vectors to facilitate the characterization and expression of the hybridizing nucleic acids in the sample. For example, such techniques may be used to isolate and clone sequences in a genomic library or cDNA library which are capable of hybridizing to the detectable probe as described in Example 30 above.

PCR primers made as described in Example 44 above may be used in forensic analyses, such as the DNA fingerprinting techniques described in Examples 46-50 below. Such analyses may utilize detectable probes or primers based on the sequences of the the 5′ ESTs or of cDNAs or genomic DNAs isolated using the 5′ ESTs.

EXAMPLE 46

Forensic Matching by DNA Sequencing

In one exemplary method, DNA samples are isolated from forensic specimens of, for example, hair, semen, blood or skin cells by conventional methods. A panel of PCR primers based on a number of the 5′ ESTs of Example 25, or cDNAs or genomic DNAs isolated

therefrom as described above, is then utilized in accordance with Example 44 to amplify DNA of approximately 100-200 bases in length from the forensic specimen. Corresponding sequences are obtained from a test subject. Each of these identification DNAs is then sequenced using standard techniques, and a simple database comparison determines the differences, if any, between the sequences from the subject and those from the sample. Statistically significant differences between the suspect's DNA sequences and those from the sample conclusively prove a lack of identity. This lack of identity can be proven, for example, with only one sequence. Identity, on the other hand, should be demonstrated with a large number of sequences, all matching. Preferably, a minimum of 50 statistically identical sequences of 100 bases in length are used to prove identity between the suspect and the sample.

EXAMPLE 47

Positive Identification by DNA Sequencing

The technique outlined in the previous example may also be used on a larger scale to provide a unique fingerprint-type identification of any individual. In this technique, primers are prepared from a large number of 5'EST sequences from Example 25, or cDNA or genomic DNA sequences obtainable therefrom. Preferably, 20 to 50 different primers are used. These primers are used to obtain a corresponding number of PCR-generated DNA segments from the individual in question in accordance with Example 44. Each of these DNA segments is sequenced, using the methods set forth in Example 46. The database of sequences generated through this procedure uniquely identifies the individual from whom the sequences were obtained. The same panel of primers may then be used at any later time to absolutely correlate tissue or other biological specimen with that individual.

EXAMPLE 48

Southern Blot Forensic Identification

The procedure of Example 47 is repeated to obtain a panel of at least 10 amplified sequences from an individual and a specimen. Preferably, the panel contains at least 50 amplified sequences. More preferably, the panel contains 100 amplified sequences. In some embodiments, the panel contains 200 amplified sequences. This PCR-generated DNA is then

digested with one or a combination of, preferably, four base specific restriction enzymes. Such enzymes are commercially available and known to those of skill in the art. After digestion, the resultant gene fragments are size separated in multiple duplicate wells on an agarose gel and transferred to nitrocellulose using Southern blotting techniques well known to those with skill in the art. For a review of Southern blotting see Davis *et al.* (Basic Methods in Molecular Biology, 1986, Elsevier Press. pp 62-65) , the disclosure of which is incorporated herein by reference..

A panel of probes based on the sequences of 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom), or fragments thereof of at least 10 bases, are radioactively or colorimetrically labeled using methods known in the art, such as nick translation or end labeling, and hybridized to the Southern blot using techniques known in the art (Davis *et al.*, supra). Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST (or cDNAs or genomic DNAs obtainable therefrom). More preferably, the probe comprises at least 20-30 consecutive nucleotides from the 5′ EST (or cDNAs or genomic DNAs obtainable therefrom). In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST (or cDNAs or genomic DNAs obtainable therefrom).

Preferably, at least 5 to 10 of these labeled probes are used, and more preferably at least about 20 or 30 are used to provide a unique pattern. The resultant bands appearing from the hybridization of a large sample of 5′ EST (or cDNAs or genomic DNAs obtainable therefrom) will be a unique identifier. Since the restriction enzyme cleavage will be different for every individual, the band pattern on the Southern blot will also be unique. Increasing the number of 5′ EST (or cDNAs or genomic DNAs obtainable therefrom) probes will provide a statistically higher level of confidence in the identification since there will be an increased number of sets of bands used for identification.

EXAMPLE 49

Dot Blot Identification Procedure

Another technique for identifying individuals using the 5′ EST sequences disclosed herein utilizes a dot blot hybridization technique.

Genomic DNA is isolated from nuclei of subject to be identified. Oligonucleotide probes of approximately 30 bp in length are synthesized that correspond to at least 10,

preferably 50 sequences from the 5′ ESTs or cDNAs or genomic DNAs obtainable therefrom. The probes are used to hybridize to the genomic DNA through conditions known to those in the art. The oligonucleotides are end labeled with P^{32} using polynucleotide kinase (Pharmacia). Dot Blots are created by spotting the genomic DNA onto nitrocellulose or the like using a vacuum dot blot manifold (BioRad, Richmond California). The nitrocellulose filter containing the genomic sequences is baked or UV linked to the filter, prehybridized and hybridized with labeled probe using techniques known in the art (Davis et *al., supra*). The ^{32}P labeled DNA fragments are sequentially hybridized with successively stringent conditions to detect minimal differences between the 30 bp sequence and the DNA. Tetramethylammonium chloride is useful for identifying clones containing small numbers of nucleotide mismatches (Wood *et al., Proc. Natl. Acad. Sci. USA* 82(6):1585-1588, 1985) which is hereby incorporated by reference. A unique pattern of dots distinguishes one individual from another individual.

5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom) or oligonucleotides containing at least 10 consecutive bases from these sequences can be used as probes in the following alternative fingerprinting technique. Preferably, the probe comprises at least 12, 15, or 17 consecutive nucleotides from the 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom). More preferably, the probe comprises at least 20-30 consecutive nucleotides from the 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom). In some embodiments, the probe comprises more than 30 nucleotides from the 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom).

Preferably, a plurality of probes having sequences from different genes are used in the alternative fingerprinting technique. Example 50 below provides a representative alternative fingerprinting procedure in which the probes are derived from 5'EST.

EXAMPLE 50

Alternative "Fingerprint" Identification Technique

20-mer oligonucleotides are prepared from a large number, e.g. 50, 100, or 200, of 5'EST using commercially available oligonucleotide services such as Genset, Paris, France. Cell samples from the test subject are processed for DNA using techniques well known to those with skill in the art. The nucleic acid is digested with restriction enzymes such as EcoRI

and XbaI: Following digestion, samples are applied to wells for electrophoresis. The procedure, as known in the art, may be modified to accommodate polyacrylamide electrophoresis, however in this example, samples containing 5 ug of DNA are loaded into wells and separated on 0.8% agarose gels. The gels are transferred onto nitrocellulose using standard Southern blotting techniques.

10 ng of each of the oligonucleotides are pooled and end-labeled with ^{32}P. The nitrocellulose is prehybridized with blocking solution and hybridized with the labeled probes. Following hybridization and washing, the nitrocellulose filter is exposed to X-Omat AR X-ray film. The resulting hybridization pattern will be unique for each individual.

It is additionally contemplated within this example that the number of probe sequences used can be varied for additional accuracy or clarity.

The proteins encoded by the extended cDNAs may also be used to generate antibodies as explained in Examples 30 and 43 in order to identify the tissue type or cell species from which a sample is derived as described in example 51.

EXAMPLE 51

Identification of Tissue Types or Cell Species by Means of Labeled Tissue Specific Antibodies

Identification of specific tissues is accomplished by the visualization of tissue specific antigens by means of antibody preparations according to Examples 30 and 43 which are conjugated, directly or indirectly to a detectable marker. Selected labeled antibody species bind to their specific antigen binding partner in tissue sections, cell suspensions, or in extracts of soluble proteins from a tissue sample to provide a pattern for qualitative or semi-qualitative interpretation.

Antisera for these procedures must have a potency exceeding that of the native preparation, and for that reason, antibodies are concentrated to a mg/ml level by isolation of the gamma globulin fraction, for example, by ion-exchange chromatography or by ammonium sulfate fractionation. Also, to provide the most specific antisera, unwanted antibodies, for example to common proteins, must be removed from the gamma globulin fraction, for example by means of insoluble immunoabsorbents, before the antibodies are

labeled with the marker. Either monoclonal or heterologous antisera is suitable for either procedure.

A. Immunohistochemical techniques

Purified, high-titer antibodies, prepared as described above, are conjugated to a detectable marker, as described, for example, by Fudenberg, Chap. 26 in: *Basic and Clinical Immunology*, 3rd Ed. Lange, Los Altos, California, 1980, or Rose,. *et al.*, Chap. 12 in: *Methods in Immunodiagnosis*, 2d Ed. John Wiley and Sons, New York (1980), the disclosures of which are incorporated herein by reference.

A fluorescent marker, either fluorescein or rhodamine, is preferred, but antibodies can also be labeled with an enzyme that supports a color producing reaction with a substrate, such as horseradish peroxidase. Markers can be added to tissue-bound antibody in a second step, as described below. Alternatively, the specific antitissue antibodies can be labeled with ferritin or other electron dense particles, and localization of the ferritin coupled antigen-antibody complexes achieved by means of an electron microscope. In yet another approach, the antibodies are radiolabeled, with, for example ^{125}I, and detected by overlaying the antibody treated preparation with photographic emulsion.

Preparations to carry out the procedures can comprise monoclonal or polyclonal antibodies to a single protein or peptide identified as specific to a tissue type, for example, brain tissue, or antibody preparations to several antigenically distinct tissue specific antigens can be used in panels, independently or in mixtures, as required.

Tissue sections and cell suspensions are prepared for immunohistochemical examination according to common histological techniques. Multiple cryostat sections (about 4 μm, unfixed) of the unknown tissue and known control, are mounted and each slide covered with different dilutions of the antibody preparation. Sections of known and unknown tissues should also be treated with preparations to provide a positive control, a negative control, for example, pre-immune sera, and a control for non-specific staining, for example, buffer.

Treated sections are incubated in a humid chamber for 30 min at room temperature, rinsed, then washed in buffer for 30-45 min. Excess fluid is blotted away, and the marker developed.

If the tissue specific antibody was not labeled in the first incubation, it can be labeled at this time in a second antibody-antibody reaction, for example, by adding fluorescein- or enzyme-conjugated antibody against the immunoglobulin class of the antiserum-producing species, for example, fluorescein labeled antibody to mouse IgG. Such labeled sera are commercially available.

The antigen found in the tissues by the above procedure can be quantified by measuring the intensity of color or fluorescence on the tissue section, and calibrating that signal using appropriate standards.

B. Identification of tissue specific soluble proteins

The visualization of tissue specific proteins and identification of unknown tissues from that procedure is carried out using the labeled antibody reagents and detection strategy as described for immunohistochemistry; however the sample is prepared according to an electrophoretic technique to distribute the proteins extracted from the tissue in an orderly array on the basis of molecular weight for detection.

A tissue sample is homogenized using a Virtis apparatus; cell suspensions are disrupted by Dounce homogenization or osmotic lysis, using detergents in either case as required to disrupt cell membranes, as is the practice in the art. Insoluble cell components such as nuclei, microsomes, and membrane fragments are removed by ultracentrifugation, and the soluble protein-containing fraction concentrated if necessary and reserved for analysis.

A sample of the soluble protein solution is resolved into individual protein species by conventional SDS polyacrylamide electrophoresis as described, for example, by Davis, *et al.*, Section 19-2 in: *Basic Methods in Molecular Biology*, Leder ed., Elsevier, New York, 1986, the disclosure of which is incorporated herein by reference, using a range of amounts of polyacrylamide in a set of gels to resolve the entire molecular weight range of proteins to be detected in the sample. A size marker is run in parallel for purposes of estimating molecular weights of the constituent proteins. Sample size for analysis is a convenient volume of from 5 to 55 μl, and containing from about 1 to 100 μg protein. An aliquot of each of the resolved proteins is transferred by blotting to a nitrocellulose filter paper, a process that maintains the pattern of resolution. Multiple copies are prepared. The procedure, known as Western Blot Analysis, is well described in Davis, L. *et al., supra* Section 19-3. One set of nitrocellulose blots is stained with Coomassie blue dye to visualize the entire set of proteins for comparison

with the antibody bound proteins. The remaining nitrocellulose filters are then incubated with a solution of one or more specific antisera to tissue specific proteins prepared as described in Examples 30 and 43. In this procedure, as in procedure A above, appropriate positive and negative sample and reagent controls are run.

In either procedure A or B, a detectable label can be attached to the primary tissue antigen-primary antibody complex according to various strategies and permutations thereof. In a straightforward approach, the primary specific antibody can be labeled; alternatively, the unlabeled complex can be bound by a labeled secondary anti-IgG antibody. In other approaches, either the primary or secondary antibody is conjugated to a biotin molecule, which can, in a subsequent step, bind an avidin conjugated marker. According to yet another strategy, enzyme labeled or radioactive protein A, which has the property of binding to any IgG, is bound in a final step to either the primary or secondary antibody.

The visualization of tissue specific antigen binding at levels above those seen in control tissues to one or more tissue specific antibodies, prepared from the gene sequences identified from extended cDNA sequences, can identify tissues of unknown origin, for example, forensic samples, or differentiated tumor tissue that has metastasized to foreign bodily sites.

In addition to their applications in forensics and identification, 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be mapped to their chromosomal locations. Example 52 below describes radiation hybrid (RH) mapping of human chromosomal regions using 5'ESTs. Example 53 below describes a representative procedure for mapping an 5′ EST to its location on a human chromosome. Example 54 below describes mapping of 5′ ESTs on metaphase chromosomes by Fluorescence In Situ Hybridization (FISH). Those skilled in the art will appreciate that the method of Examples 52-54 may also be used to map cDNAs or genomic DNAs obtainable from the 5′ ESTs to their chromosomal locations.

<u>2. Use of 5' ESTs or Sequences Obtainable Therefrom or Portions Thereof in Chromosome Mapping</u>

EXAMPLE 52

Radiation hybrid mapping of 5'ESTs to the human genome

Radiation hybrid (RH) mapping is a somatic cell genetic approach that can be used for high resolution mapping of the human genome. In this approach, cell lines containing one or more human chromosomes are lethally irradiated, breaking each chromosome into fragments whose size depends on the radiation dose. These fragments are rescued by fusion with cultured rodent cells, yielding subclones containing different portions of the human genome. This technique is described by Benham *et al.*, *Genomics* 4:509-517, 1989; and Cox *et al.*, *Science* 250:245-250, 1990, the entire contents of which are hereby incorporated by reference. The random and independent nature of the subclones permits efficient mapping of any human genome marker. Human DNA isolated from a panel of 80-100 cell lines provides a mapping reagent for ordering 5'EST. In this approach, the frequency of breakage between markers is used to measure distance, allowing construction of fine resolution maps as has been done using conventional ESTs (Schuler *et al.*, *Science* 274:540-546, 1996, hereby incorporated by reference).

RH mapping has been used to generate a high-resolution whole genome radiation hybrid map of human chromosome 17q22-q25.3 across the genes for growth hormone (GH) and thymidine kinase (TK) (Foster *et al.*, *Genomics* 33:185-192, 1996), the region surrounding the Gorlin syndrome gene (Obermayr *et al.*, *Eur. J. Hum. Genet.* 4:242-245, 1996), 60 loci covering the entire short arm of chromosome 12 (Raeymaekers *et al.*, *Genomics* 29:170-178, 1995), the region of human chromosome 22 containing the neurofibromatosis type 2 locus (Frazer *et al.*, *Genomics* 14:574-584, 1992) and 13 loci on the long arm of chromosome 5 (Warrington *et al.*, *Genomics* 11:701-708, 1991).

EXAMPLE 53

Mapping of 5'ESTs to HumanChromosomes using PCR techniques

5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) may be assigned to human chromosomes using PCR based methodologies. In such approaches, oligonucleotide primer pairs are designed from the 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) to minimize the chance of amplifying through an intron. Preferably, the oligonucleotide primers are 18-23 bp in length and are designed for PCR amplification. The

creation of PCR primers from known sequences is well known to those with skill in the art. For a review of PCR technology see Erlich in *PCR Technology; Principles and Applications for DNA Amplification*, Freeman and Co., New York, 1992, the disclosure of which is incorporated herein by reference..

The primers are used in polymerase chain reactions (PCR) to amplify templates from total human genomic DNA. PCR conditions are as follows: 60 ng of genomic DNA is used as a template for PCR with 80 ng of each oligonucleotide primer, 0.6 unit of Taq polymerase, and 1 μCu of a ^{32}P-labeled deoxycytidine triphosphate. The PCR is performed in a microplate thermocycler (Techne) under the following conditions: 30 cycles of 94°C, 1.4 min; 55°C, 2 min; and 72°C, 2 min; with a final extension at 72°C for 10 min. The amplified products are analyzed on a 6% polyacrylamide sequencing gel and visualized by autoradiography. If the length of the resulting PCR product is identical to the distance between the ends of the primer sequences in the extended cDNA from which the primers are derived, then the PCR reaction is repeated with DNA templates from two panels of human-rodent somatic cell hybrids, BIOS PCRable DNA (BIOS Corporation) and NIGMS Human-Rodent Somatic Cell Hybrid Mapping Panel Number 1 (NIGMS, Camden, NJ).

PCR is used to screen a series of somatic cell hybrid cell lines containing defined sets of human chromosomes for the presence of a given 5′ EST (or cDNA or genomic DNA obtainable therefrom). DNA is isolated from the somatic hybrids and used as starting templates for PCR reactions using the primer pairs from the 5′ EST (or cDNA or genomic DNA obtainable therefrom). Only those somatic cell hybrids with chromosomes containing the human gene corresponding to the 5′ EST (or cDNA or genomic DNA obtainable therefrom) will yield an amplified fragment. The 5′ EST (or cDNA or genomic DNA obtainable therefrom) are assigned to a chromosome by analysis of the segregation pattern of PCR products from the somatic hybrid DNA templates. The single human chromosome present in all cell hybrids that give rise to an amplified fragment is the chromosome containing that 5′EST (or cDNA or genomic DNA obtainable therefrom). For a review of techniques and analysis of results from somatic cell gene mapping experiments, see Ledbetter *et al.*, *Genomics* 6:475-481, 1990, the disclosure of which is incorporated herein by reference.

EXAMPLE 54

Mapping of Extended 5' ESTs to Chromosomes Using Fluorescence *In Situ* Hybridization

Fluorescence in situ hybridization allows the 5'EST (or cDNA or genomic DNA obtainable therefrom) to be mapped to a particular location on a given chromosome. The chromosomes to be used for fluorescence in situ hybridization techniques may be obtained from a variety of sources including cell cultures, tissues, or whole blood.

In a preferred embodiment, chromosomal localization of an 5'EST (or cDNA or genomic DNA obtainable therefrom) is obtained by FISH as described by Cherif *et al.* (*Proc. Natl. Acad. Sci. U.S.A.*, 87:6639-6643, 1990) , the disclosure of which is incorporated herein by reference.. Metaphase chromosomes are prepared from phytohemagglutinin (PHA)-stimulated blood cell donors. PHA-stimulated lymphocytes from healthy males are cultured for 72 h in RPMI-1640 medium. For synchronization, methotrexate (10 μM) is added for 17 h, followed by addition of 5-bromodeoxyuridine (5-BrdU, 0.1 mM) for 6 h. Colcemid (1 μg/ml) is added for the last 15 min before harvesting the cells. Cells are collected, washed in RPMI, incubated with a hypotonic solution of KCl (75 mM) at 37°C for 15 min and fixed in three changes of methanol:acetic acid (3:1). The cell suspension is dropped onto a glass slide and air dried. The 5'EST (or cDNA or genomic DNA obtainable therefrom) is labeled with biotin-16 dUTP by nick translation according to the manufacturer's instructions (Bethesda Research Laboratories, Bethesda, MD), purified using a Sephadex G-50 column (Pharmacia, Upsala, Sweden) and precipitated. Just prior to hybridization, the DNA pellet is dissolved in hybridization buffer (50% formamide, 2 X SSC, 10% dextran sulfate, 1 mg/ml sonicated salmon sperm DNA, pH 7) and the probe is denatured at 70°C for 5-10 min.

Slides kept at -20°C are treated for 1 h at 37°C with RNase A (100 μg/ml), rinsed three times in 2 X SSC and dehydrated in an ethanol series. Chromosome preparations are denatured in 70% formamide, 2 X SSC for 2 min at 70°C, then dehydrated at 4°C. The slides are treated with proteinase K (10 μg/100 ml in 20 mM Tris-HCl, 2 mM $CaCl_2$) at 37°C for 8 min and dehydrated. The hybridization mixture containing the probe is placed on the slide, covered with a coverslip, sealed with rubber cement and incubated overnight in a humid chamber at 37°C. After hybridization and post-hybridization washes, the biotinylated probe is detected by avidin-FITC and amplified with additional layers of biotinylated goat anti-avidin

and avidin-FITC. For chromosomal localization, fluorescent R-bands are obtained as previously described (Cherif *et al.*, *supra*.). The slides are observed under a LEICA fluorescence microscope (DMRXA). Chromosomes are counterstained with propidium iodide and the fluorescent signal of the probe appears as two symmetrical yellow-green spots on both chromatids of the fluorescent R-band chromosome (red). Thus, a particular 5'EST (or cDNA or genomic DNA obtainable therefrom) may be localized to a particular cytogenetic R-band on a given chromosome.

Once the 5'EST (or cDNA or genomic DNA obtainable therefrom) have been assigned to particular chromosomes using the techniques described in Examples 52-54 above, they may be utilized to construct a high resolution map of the chromosomes on which they are located or to identify the chromosomes in a sample.

EXAMPLE 55

Use of 5'EST to Construct or Expand Chromosome Maps

Chromosome mapping involves assigning a given unique sequence to a particular chromosome as described above. Once the unique sequence has been mapped to a given chromosome, it is ordered relative to other unique sequences located on the same chromosome. One approach to chromosome mapping utilizes a series of yeast artificial chromosomes (YACs) bearing several thousand long inserts derived from the chromosomes of the organism from which the extended cDNAs (or genomic DNAs obtainable therefrom) are obtained. This approach is described in Nagaraja *et al.*, *Genome Research* 7:210-222, 1997, the disclosure of which is incorporated herein by reference. Briefly, in this approach each chromosome is broken into overlapping pieces which are inserted into the YAC vector. The YAC inserts are screened using PCR or other methods to determine whether they include the 5'EST (or cDNA or genomic DNA obtainable therefrom) whose position is to be determined. Once an insert has been found which includes the 5'EST (or cDNA or genomic DNA obtainable therefrom), the insert can be analyzed by PCR or other methods to determine whether the insert also contains other sequences known to be on the chromosome or in the region from which the 5'EST (or cDNA or genomic DNA obtainable therefrom) was derived. This process can be repeated for each insert in the YAC library to determine the

location of each of the extended cDNAs (or genomic DNAs obtainable therefrom) relative to one another and to other known chromosomal markers. In this way, a high resolution map of the distribution of numerous unique markers along each of the organisms chromosomes may be obtained.

As described in Example 56 below extended cDNAs (or genomic DNAs obtainable therefrom) may also be used to identify genes associated with a particular phenotype, such as hereditary disease or drug response.

3. Use of 5'ESTs or Sequences Obtained Therefrom or Fragments Thereof in Gene Identification

EXAMPLE 56

Identification of genes associated with hereditary diseases or drug response

This example illustrates an approach useful for the association of 5'ESTs (or cDNA or genomic DNA obtainable therefrom) with particular phenotypic characteristics. In this example, a particular 5'EST (or cDNA or genomic DNA obtainable therefrom) is used as a test probe to associate that 5'EST (or cDNA or genomic DNA obtainable therefrom) with a particular phenotypic characteristic.

5'ESTs (or cDNA or genomic DNA obtainable therefrom) are mapped to a particular location on a human chromosome using techniques such as those described in Examples 52 and 53 or other techniques known in the art. A search of Mendelian Inheritance in Man (McKusick in *Mendelian Inheritance in Man* (available on line through Johns Hopkins University Welch Medical Library) reveals the region of the human chromosome which contains the 5'EST (or cDNA or genomic DNA obtainable therefrom) to be a very gene rich region containing several known genes and several diseases or phenotypes for which genes have not been identified. The gene corresponding to this 5'EST (or cDNA or genomic DNA obtainable therefrom) thus becomes an immediate candidate for each of these genetic diseases.

Cells from patients with these diseases or phenotypes are isolated and expanded in culture. PCR primers from the 5'EST (or cDNA or genomic DNA obtainable

therefrom) are used to screen genomic DNA, mRNA or cDNA obtained from the patients. 5'ESTs (or cDNA or genomic DNA obtainable therefrom) that are not amplified in the patients can be positively associated with a particular disease by further analysis. Alternatively, the PCR analysis may yield fragments of different lengths when the samples are derived from an individual having the phenotype associated with the disease than when the sample is derived from a healthy individual, indicating that the gene containing the 5'EST may be responsible for the genetic disease.

VI. Use of 5'EST (or cDNA or Genomic DNA Obtainable Therefrom) to Construct Vectors

The present 5'ESTs (or cDNA or genomic DNA obtainable therefrom) may also be used to construct secretion vectors capable of directing the secretion of the proteins encoded by genes therein. Such secretion vectors may facilitate the purification or enrichment of the proteins encoded by genes inserted therein by reducing the number of background proteins from which the desired protein must be purified or enriched. Exemplary secretion vectors are described in Example 57 below.

1. Construction of Secretion Vectors

EXAMPLE 57

Construction of Secretion Vectors

The secretion vectors include a promoter capable of directing gene expression in the host cell, tissue, or organism of interest. Such promoters include the Rous Sarcoma Virus promoter, the SV40 promoter, the human cytomegalovirus promoter, and other promoters familiar to those skilled in the art.

A signal sequence from a 5' EST (or cDNAs or genomic DNAs obtainable therefrom) is operably linked to the promoter such that the mRNA transcribed from the promoter will direct the translation of the signal peptide. The host cell, tissue, or organism may be any cell, tissue, or organism which recognizes the signal peptide encoded by the signal sequence in the 5' EST (or cDNA or genomic DNA obtainable therefrom). Suitable hosts include mammalian cells, tissues or organisms, avian cells, tissues, or organisms, insect cells, tissues or organisms, or yeast.

In addition, the secretion vector contains cloning sites for inserting genes encoding the proteins which are to be secreted. The cloning sites facilitate the cloning of the insert gene in frame with the signal sequence such that a fusion protein in which the signal peptide is fused to the protein encoded by the inserted gene is expressed from the mRNA transcribed from the promoter. The signal peptide directs the extracellular secretion of the fusion protein.

The secretion vector may be DNA or RNA and may integrate into the chromosome of the host, be stably maintained as an extrachromosomal replicon in the host, be an artificial chromosome, or be transiently present in the host. Many nucleic acid backbones suitable for use as secretion vectors are known to those skilled in the art, including retroviral vectors, SV40 vectors, Bovine Papilloma Virus vectors, yeast integrating plasmids, yeast episomal plasmids, yeast artificial chromosomes, human artificial chromosomes, P element vectors, baculovirus vectors, or bacterial plasmids capable of being transiently introduced into the host.

The secretion vector may also contain a polyA signal such that the polyA signal is located downstream of the gene inserted into the secretion vector.

After the gene encoding the protein for which secretion is desired is inserted into the secretion vector, the secretion vector is introduced into the host cell, tissue, or organism using calcium phosphate precipitation, DEAE-Dextran, electroporation, liposome-mediated transfection, viral particles or as naked DNA. The protein encoded by the inserted gene is then purified or enriched from the supernatant using conventional techniques such as ammonium sulfate precipitation, immunoprecipitation, immunochromatography, size exclusion chromatography, ion exchange chromatography, and HPLC. Alternatively, the secreted protein may be in a sufficiently enriched or pure state in the supernatant or growth media of the host to permit it to be used for its intended purpose without further enrichment.

The signal sequences may also be inserted into vectors designed for gene therapy. In such vectors, the signal sequence is operably linked to a promoter such that mRNA transcribed from the promoter encodes the signal peptide. A cloning site is located downstream of the signal sequence such that a gene encoding a protein whose secretion is desired may readily be inserted into the vector and fused to the signal sequence. The vector is introduced into an appropriate host cell. The protein expressed from the promoter is secreted extracellularly, thereby producing a therapeutic effect.

The 5′ ESTs may also be used to clone sequences located upstream of the 5′ ESTs which are capable of regulating gene expression, including promoter sequences, enhancer sequences, and other upstream sequences which influence transcription or translation levels. Once identified and cloned, these upstream regulatory sequences may be used in expression vectors designed to direct the expression of an inserted gene in a desired spatial, temporal, developmental, or quantitative fashion. Example 58 describes a method for cloning sequences upstream of the extended cDNAs or 5′ ESTs.

2. Identification of Upstream Sequences With Promoting or Regulatory Activities

EXAMPLE 58

Use of Extended cDNAs or 5′ ESTs to Clone Upstream Sequences from Genomic DNA

Sequences derived from extended cDNAs or 5′ ESTs may be used to isolate the promoters of the corresponding genes using chromosome walking techniques. In one chromosome walking technique, which utilizes the GenomeWalker™ kit available from Clontech, five complete genomic DNA samples are each digested with a different restriction enzyme which has a 6 base recognition site and leaves a blunt end. Following digestion, oligonucleotide adapters are ligated to each end of the resulting genomic DNA fragments.

For each of the five genomic DNA libraries, a first PCR reaction is performed according to the manufacturer's instructions (which are incorporated herein by reference) using an outer adaptor primer provided in the kit and an outer gene specific primer. The gene specific primer should be selected to be specific for the extended cDNA or 5′ EST of interest and should have a melting temperature, length, and location in the extended cDNA or 5'EST which is consistent with its use in PCR reactions. Each first PCR reaction contains 5 ng of genomic DNA, 5 μl of 10X Tth reaction buffer, 0.2 mM of each dNTP, 0.2 μM each of outer adaptor primer and outer gene specific primer, 1.1 mM of $Mg(OAc)_2$, and 1 μl of the Tth polymerase 50X mix in a total volume of 50 μl. The reaction cycle for the first PCR reaction is as follows: 1 min - 94°C / 2 sec - 94°C, 3 min - 72°C (7 cycles) / 2 sec - 94°C, 3 min - 67°C (32 cycles) / 5 min - 67°C.

The product of the first PCR reaction is diluted and used as a template for a second PCR reaction according to the manufacturer's instructions using a pair of nested

primers which are located internally on the amplicon resulting from the first PCR reaction. For example, 5 µl of the reaction product of the first PCR reaction mixture may be diluted 180 times. Reactions are made in a 50 µl volume having a composition identical to that of the first PCR reaction except the nested primers are used. The first nested primer is specific for the adaptor, and is provided with the GenomeWalker™ kit. The second nested primer is specific for the particular extended cDNA or 5′ EST for which the promoter is to be cloned and should have a melting temperature, length, and location in the extended cDNA or 5′ EST which is consistent with its use in PCR reactions. The reaction parameters of the second PCR reaction are as follows: 1 min - 94°C / 2 sec - 94°C, 3 min - 72°C (6 cycles) / 2 sec - 94°C, 3 min - 67°C (25 cycles) / 5 min - 67°C. The product of the second PCR reaction is purified, cloned, and sequenced using standard techniques.

Alternatively, two or more human genomic DNA libraries can be constructed by using two or more restriction enzymes. The digested genomic DNA is cloned into vectors which can be converted into single stranded, circular, or linear DNA. A biotinylated oligonucleotide comprising at least 15 nucleotides from the extended cDNA or 5′ EST sequence is hybridized to the single stranded DNA. Hybrids between the biotinylated oligonucleotide and the single stranded DNA containing the extended cDNA or EST sequence are isolated as described in Example 29 above. Thereafter, the single stranded DNA containing the extended cDNA or EST sequence is released from the beads and converted into double stranded DNA using a primer specific for the extended cDNA or 5′ EST sequence or a primer corresponding to a sequence included in the cloning vector. The resulting double stranded DNA is transformed into bacteria. DNAs containing the 5′ EST or extended cDNA sequences are identified by colony PCR or colony hybridization.

Once the upstream genomic sequences have been cloned and sequenced as described above, prospective promoters and transcription start sites within the upstream sequences may be identified by comparing the sequences upstream of the extended cDNAs or 5′ ESTs with databases containing known transcription start sites, transcription factor binding sites, or promoter sequences.

In addition, promoters in the upstream sequences may be identified using promoter reporter vectors as described in Example .

EXAMPLE 59

Identification of Promoters in Cloned Upstream Sequences

The genomic sequences upstream of the extended cDNAs or 5′ ESTs are cloned into a suitable promoter reporter vector, such as the pSEAP-Basic, pSEAP-Enhancer, pβgal-Basic, pβgal-Enhancer, or pEGFP-1 Promoter Reporter vectors available from Clontech. Briefly, each of these promoter reporter vectors include multiple cloning sites positioned upstream of a reporter gene encoding a readily assayable protein such as secreted alkaline phosphatase, β galactosidase, or green fluorescent protein. The sequences upstream of the extended cDNAs or 5′ ESTs are inserted into the cloning sites upstream of the reporter gene in both orientations and introduced into an appropriate host cell. The level of reporter protein is assayed and compared to the level obtained from a vector which lacks an insert in the cloning site. The presence of an elevated expression level in the vector containing the insert with respect to the control vector indicates the presence of a promoter in the insert. If necessary, the upstream sequences can be cloned into vectors which contain an enhancer for augmenting transcription levels from weak promoter sequences. A significant level of expression above that observed with the vector lacking an insert indicates that a promoter sequence is present in the inserted upstream sequence.

Appropriate host cells for the promoter reporter vectors may be chosen based on the results of the above described determination of expression patterns of the extended cDNAs and ESTs. For example, if the expression pattern analysis indicates that the mRNA corresponding to a particular extended cDNA or 5′ EST is expressed in fibroblasts, the promoter reporter vector may be introduced into a human fibroblast cell line.

Promoter sequences within the upstream genomic DNA may be further defined by constructing nested deletions in the upstream DNA using conventional techniques such as Exonuclease III digestion. The resulting deletion fragments can be inserted into the promoter reporter vector to determine whether the deletion has reduced or obliterated promoter activity. In this way, the boundaries of the promoters may be defined. If desired, potential individual regulatory sites within the promoter may be identified using site directed

mutagenesis or linker scanning to obliterate potential transcription factor binding sites within the promoter individually or in combination. The effects of these mutations on transcription levels may be determined by inserting the mutations into the cloning sites in the promoter reporter vectors.

EXAMPLE 60

Cloning and Identification of Promoters

Using the method described in Example 58 above with 5' ESTs, sequences upstream of several genes were obtained. Using the primer pairs GGG AAG ATG GAG ATA GTA TTG CCT G (SEQ ID NO:29) and CTG CCA TGT ACA TGA TAG AGA GAT TC (SEQ ID NO:30), the promoter having the internal designation P13H2 (SEQ ID NO:31) was obtained.

Using the primer pairs GTA CCA GGGG ACT GTG ACC ATT GC (SEQ ID NO:32) and CTG TGA CCA TTG CTC CCA AGA GAG (SEQ ID NO:33), the promoter having the internal designation P15B4 (SEQ ID NO:34) was obtained.

Using the primer pairs CTG GGA TGG AAG GCA CGG TA (SEQ ID NO:35) and GAG ACC ACA CAG CTA GAC AA (SEQ ID NO:36), the promoter having the internal designation P29B6 (SEQ ID NO:37) was obtained.

Figure 4 provides a schematic description of the promoters isolated and the way they are assembled with the corresponding 5' tags. The upstream sequences were screened for the presence of motifs resembling transcription factor binding sites or known transcription start sites using the computer program MatInspector release 2.0, August 1996.

Table VII describes the transcription factor binding sites present in each of these promoters. The columns labeled matrice provides the name of the MatInspector matrix used. The column labeled position provides the 5' position of the promoter site. Numeration of the sequence starts from the transcription site as determined by matching the genomic sequence with the 5' EST sequence. The column labeled "orientation" indicates the DNA strand on which the site is found, with the + strand being the coding strand as determined by matching the genomic sequence with the sequence of the 5' EST. The column labeled "score" provides the MatInspector score found for this site. The column labeled "length" provides the length

of the site in nucleotides. The column labeled "sequence" provides the sequence of the site found.

Bacterial clones containing plasmids containing the promoter sequences described above described above are presently stored in the inventor's laboratories under the internal identification numbers provided above. The inserts may be recovered from the deposited materials by growing an aliquot of the appropriate bacterial clone in the appropriate medium. The plasmid DNA can then be isolated using plasmid isolation procedures familiar to those skilled in the art such as alkaline lysis minipreps or large scale alkaline lysis plasmid isolation procedures. If desired the plasmid DNA may be further enriched by centrifugation on a cesium chloride gradient, size exclusion chromatography, or anion exchange chromatography. The plasmid DNA obtained using these procedures may then be manipulated using standard cloning techniques familiar to those skilled in the art. Alternatively, a PCR can be done with primers designed at both ends of the EST insertion. The PCR product which corresponds to the 5′ EST can then be manipulated using standard cloning techniques familiar to those skilled in the art.

The promoters and other regulatory sequences located upstream of the extended cDNAs or 5′ ESTs may be used to design expression vectors capable of directing the expression of an inserted gene in a desired spatial, temporal, developmental, or quantitative manner. A promoter capable of directing the desired spatial, temporal, developmental, and quantitative patterns may be selected using the results of the expression analysis described in Example 26 above. For example, if a promoter which confers a high level of expression in muscle is desired, the promoter sequence upstream of an extended cDNA or 5′ EST derived from an mRNA which is expressed at a high level in muscle, as determined by the method of Example 26, may be used in the expression vector.

Preferably, the desired promoter is placed near multiple restriction sites to facilitate the cloning of the desired insert downstream of the promoter, such that the promoter is able to drive expression of the inserted gene. The promoter may be inserted in conventional nucleic acid backbones designed for extrachromosomal replication, integration into the host chromosomes or transient expression. Suitable backbones for the present expression vectors include retroviral backbones, backbones from eukaryotic episomes such as SV40 or Bovine Papilloma Virus, backbones from bacterial episomes, or artificial chromosomes.

Preferably, the expression vectors also include a polyA signal downstream of the multiple restriction sites for directing the polyadenylation of mRNA transcribed from the gene inserted into the expression vector.

Following the identification of promoter sequences using the procedures of Examples 58-60, proteins which interact with the promoter may be identified as described in Example 61 below.

EXAMPLE 61

Identification of Proteins Which Interact with Promoter Sequences, Upstream Regulatory Sequences, or mRNA

Sequences within the promoter region which are likely to bind transcription factors may be identified by homology to known transcription factor binding sites or through conventional mutagenesis or deletion analyses of reporter plasmids containing the promoter sequence. For example, deletions may be made in a reporter plasmid containing the promoter sequence of interest operably linked to an assayable reporter gene. The reporter plasmids carrying various deletions within the promoter region are transfected into an appropriate host cell and the effects of the deletions on expression levels is assessed. Transcription factor binding sites within the regions in which deletions reduce expression levels may be further localized using site directed mutagenesis, linker scanning analysis, or other techniques familiar to those skilled in the art.

Nucleic acids encoding proteins which interact with sequences in the promoter may be identified using one-hybrid systems such as those described in the manual accompanying the Matchmaker One-Hybrid System kit available from Clontech (Catalog No. K1603-1), the disclosure of which is incorporated herein by reference. Briefly, the Matchmaker One-hybrid system is used as follows. The target sequence for which it is desired to identify binding proteins is cloned upstream of a selectable reporter gene and integrated into the yeast genome. Preferably, multiple copies of the target sequences are inserted into the reporter plasmid in tandem. A library comprised of fusions between cDNAs to be evaluated for the ability to bind to the promoter and the activation domain of a yeast transcription factor, such as GAL4, is transformed into the yeast strain containing the integrated reporter sequence. The yeast are plated on selective media to

select cells expressing the selectable marker linked to the promoter sequence. The colonies which grow on the selective media contain genes encoding proteins which bind the target sequence. The inserts in the genes encoding the fusion proteins are further characterized by sequencing. In addition, the inserts may be inserted into expression vectors or *in vitro* transcription vectors. Binding of the polypeptides encoded by the inserts to the promoter DNA may be confirmed by techniques familiar to those skilled in the art, such as gel shift analysis or DNAse protection analysis.

VII. Use of 5' ESTs (or cDNAs or Genomic DNAs Obtainable Therefrom) in Gene Therapy

The present invention also comprises the use of 5'ESTs (or cDNA or genomic DNA obtainable therefrom) in gene therapy strategies, including antisense and triple helix strategies as described in Examples 62 and 63 below. In antisense approaches, nucleic acid sequences complementary to an mRNA are hybridized to the mRNA intracellularly, thereby blocking the expression of the protein encoded by the mRNA. The antisense sequences may prevent gene expression through a variety of mechanisms. For example, the antisense sequences may inhibit the ability of ribosomes to translate the mRNA. Alternatively, the antisense sequences may block transport of the mRNA from the nucleus to the cytoplasm, thereby limiting the amount of mRNA available for translation. Another mechanism through which antisense sequences may inhibit gene expression is by interfering with mRNA splicing. In yet another strategy, the antisense nucleic acid may be incorporated in a ribozyme capable of specifically cleaving the target mRNA.

EXAMPLE 62

Preparation and Use of Antisense Oligonucleotides

The antisense nucleic acid molecules to be used in gene therapy may be either DNA or RNA sequences. They may comprise a sequence complementary to the sequence of the 5'EST (or cDNA or genomic DNA obtainable therefrom). The antisense nucleic acids should have a length and melting temperature sufficient to permit formation of an intracellular duplex with sufficient stability to inhibit the expression of the mRNA in the duplex. Strategies for designing antisense nucleic acids suitable for use in gene therapy are disclosed in Green *et*

al., *Ann. Rev. Biochem.* 55:569-597, 1986; and Izant and Weintraub, *Cell* 36:1007-1015, 1984, which are hereby incorporated by reference.

In some strategies, antisense molecules are obtained from a nucleotide sequence encoding a protein by reversing the orientation of the coding region with respect to a promoter so as to transcribe the opposite strand from that which is normally transcribed in the cell. The antisense molecules may be transcribed using *in vitro* transcription systems such as those which employ T7 or SP6 polymerase to generate the transcript. Another approach involves transcription of the antisense nucleic acids *in vivo* by operably linking DNA containing the antisense sequence to a promoter in an expression vector.

Alternatively, oligonucleotides which are complementary to the strand normally transcribed in the cell may be synthesized *in vitro*. Thus, the antisense nucleic acids are complementary to the corresponding mRNA and are capable of hybridizing to the mRNA to create a duplex. In some embodiments, the antisense sequences may contain modified sugar phosphate backbones to increase stability and make them less sensitive to RNase activity. Examples of modifications suitable for use in antisense strategies are described by Rossi *et al.*, *Pharmacol. Ther.* 50(2):245-254, 1991, which is hereby incorporated by reference.

Various types of antisense oligonucleotides complementary to the sequence of the 5'EST (or cDNA or genomic DNA obtainable therefrom) may be used. In one preferred embodiment, stable and semi-stable antisense oligonucleotides described in International Application No. PCT WO94/23026, hereby incorporated by reference, are used. In these molecules, the 3' end or both the 3' and 5' ends are engaged in intramolecular hydrogen bonding between complementary base pairs. These molecules are better able to withstand exonuclease attacks and exhibit increased stability compared to conventional antisense oligonucleotides.

In another preferred embodiment, the antisense oligodeoxynucleotides against herpes simplex virus types 1 and 2 described in International Application No. WO 95/04141, hereby incorporated by reference, are used.

In yet another preferred embodiment, the covalently cross-linked antisense oligonucleotides described in International Application No. WO 96/31523, hereby incorporated by reference, are used. These double- or single-stranded oligonucleotides comprise one or more, respectively, inter- or intra-oligonucleotide covalent cross-linkages,

wherein the linkage consists of an amide bond between a primary amine group of one strand and a carboxyl group of the other strand or of the same strand, respectively, the primary amine group being directly substituted in the 2' position of the strand nucleotide monosaccharide ring, and the carboxyl group being carried by an aliphatic spacer group substituted on a nucleotide or nucleotide analog of the other strand or the same strand, respectively.

The antisense oligodeoxynucleotides and oligonucleotides disclosed in International Application No. WO 92/18522, incorporated by reference, may also be used. These molecules are stable to degradation and contain at least one transcription control recognition sequence which binds to control proteins and are effective as decoys therefore. These molecules may contain "hairpin" structures, "dumbbell" structures, "modified dumbbell" structures, "cross-linked" decoy structures and "loop" structures.

In another preferred embodiment, the cyclic double-stranded oligonucleotides described in European Patent Application No. 0 572 287 A2, hereby incorporated by reference are used. These ligated oligonucleotide "dumbbells" contain the binding site for a transcription factor and inhibit expression of the gene under control of the transcription factor by sequestering the factor.

Use of the closed antisense oligonucleotides disclosed in International Application No. WO 92/19732, hereby incorporated by reference, is also contemplated. Because these molecules have no free ends, they are more resistant to degradation by exonucleases than are conventional oligonucleotides. These oligonucleotides may be multifunctional, interacting with several regions which are not adjacent to the target mRNA.

The appropriate level of antisense nucleic acids required to inhibit gene expression may be determined using *in vitro* expression analysis. The antisense molecule may be introduced into the cells by diffusion, injection, infection, transfection or h-region-mediated import using procedures known in the art. For example, the antisense nucleic acids can be introduced into the body as a bare or naked oligonucleotide, oligonucleotide encapsulated in lipid, oligonucleotide sequence encapsidated by viral protein, or as an oligonucleotide operably linked to a promoter contained in an expression vector. The expression vector may be any of a variety of expression vectors known in the art, including retroviral or viral vectors,

vectors capable of extrachromosomal replication, or integrating vectors. The vectors may be DNA or RNA.

The antisense molecules are introduced onto cell samples at a number of different concentrations preferably between $1x10^{-10}$M to $1x10^{-4}$M. Once the minimum concentration that can adequately control gene expression is identified, the optimized dose is translated into a dosage suitable for use *in vivo*. For example, an inhibiting concentration in culture of $1x10^{-7}$ translates into a dose of approximately 0.6 mg/kg bodyweight. Levels of oligonucleotide approaching 100 mg/kg bodyweight or higher may be possible after testing the toxicity of the oligonucleotide in laboratory animals. It is additionally contemplated that cells from the vertebrate are removed, treated with the antisense oligonucleotide, and reintroduced into the vertebrate.

It is further contemplated that the antisense oligonucleotide sequence is incorporated into a ribozyme sequence to enable the antisense to specifically bind and cleave its target mRNA. For technical applications of ribozyme and antisense oligonucleotides see Rossi *et al.*, *supra*.

In a preferred application of this invention, the polypeptide encoded by the gene is first identified, so that the effectiveness of antisense inhibition on translation can be monitored using techniques that include but are not limited to antibody-mediated tests such as RIAs and ELISA, functional assays, or radiolabeling.

The 5′ ESTs of the present invention (or cDNAs or genomic DNAs obtainable therefrom) may also be used in gene therapy approaches based on intracellular triple helix formation. Triple helix oligonucleotides are used to inhibit transcription from a genome. They are particularly useful for studying alterations in cell activity as it is associated with a particular gene. The 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom) of the present invention or, more preferably, a portion of those sequences, can be used to inhibit gene expression in individuals having diseases associated with expression of a particular gene. Similarly, a portion of 5′ EST sequences (or cDNAs or genomic DNAs obtainable therefrom) can be used to study the effect of inhibiting transcription of a particular gene within a cell. Traditionally, homopurine sequences were considered the most useful for triple helix strategies. However, homopyrimidine sequences can also inhibit gene expression. Such homopyrimidine oligonucleotides bind to the major groove at

homopurine:homopyrimidine sequences. Thus, both types of sequences from the 5'EST or from the gene corresponding to the 5'EST are contemplated within the scope of this invention.

EXAMPLE 63

Preparation and Use of Triple Helix Probes

The sequences of the 5′ ESTs (or cDNAs or genomic DNAs obtainable therefrom) are scanned to identify 10-mer to 20-mer homopyrimidine or homopurine stretches which could be used in triple-helix based strategies for inhibiting gene expression. Following identification of candidate homopyrimidine or homopurine stretches, their efficiency in inhibiting gene expression is assessed by introducing varying amounts of oligonucleotides containing the candidate sequences into tissue culture cells which normally express the target gene. The oligonucleotides may be prepared on an oligonucleotide synthesizer or they may be purchased commercially from a company specializing in custom oligonucleotide synthesis, such as GENSET, Paris, France.

The oligonucleotides may be introduced into the cells using a variety of methods known to those skilled in the art, including but not limited to calcium phosphate precipitation, DEAE-Dextran, electroporation, liposome-mediated transfection or native uptake.

Treated cells are monitored for altered cell function or reduced gene expression using techniques such as Northern blotting, RNase protection assays, or PCR based strategies to monitor the transcription levels of the target gene in cells which have been treated with the oligonucleotide. The cell functions to be monitored are predicted based upon the homologies of the target gene corresponding to the extended cDNA from which the oligonucleotide was derived with known gene sequences that have been associated with a particular function. The cell functions can also be predicted based on the presence of abnormal physiologies within cells derived from individuals with a particular inherited disease, particularly when the extended cDNA is associated with the disease using techniques described in Example 56.

The oligonucleotides which are effective in inhibiting gene expression in tissue culture cells may then be introduced *in vivo* using the techniques described above and in Example 62 at a dosage calculated based on the *in vitro* results, as described in Example 62.

In some embodiments, the natural (beta) anomers of the oligonucleotide units can be replaced with alpha anomers to render the oligonucleotide more resistant to nucleases. Further, an intercalating agent such as ethidium bromide, or the like, can be attached to the 3′ end of the alpha oligonucleotide to stabilize the triple helix. For information on the generation of oligonucleotides suitable for triple helix formation see Griffin *et al.*, *Science* 245:967-971, 1989, which is hereby incorporated by this reference.

EXAMPLE 64

Use of cDNAs Obtained Using the 5′ ESTs to Express an Encoded Protein in a Host Organism

The cDNAs obtained as described above using the 5′ ESTs of the present invention may also be used to express an encoded protein in a host organism to produce a beneficial effect. In such procedures, the encoded protein may be transiently expressed in the host organism or stably expressed in the host organism. The encoded protein may have any of the activities described above. The encoded protein may be a protein which the host organism lacks or, alternatively, the encoded protein may augment the existing levels of the protein in the host organism.

A full length extended cDNA encoding the signal peptide and the mature protein, or an extended cDNA encoding only the mature protein is introduced into the host organism. The extended cDNA may be introduced into the host organism using a variety of techniques known to those of skill in the art. For example, the extended cDNA may be injected into the host organism as naked DNA such that the encoded protein is expressed in the host organism, thereby producing a beneficial effect.

Alternatively, the extended cDNA may be cloned into an expression vector downstream of a promoter which is active in the host organism. The expression vector may be any of the expression vectors designed for use in gene therapy, including viral or retroviral vectors. The expression vector may be directly introduced into the host organism such that the encoded protein is expressed in the host organism to produce a beneficial effect. In another approach, the expression vector may be introduced into cells *in vitro*. Cells containing the expression vector are thereafter selected and introduced into the host organism, where they express the encoded protein to produce a beneficial effect.

EXAMPLE 65

Use of Signal Peptides Encoded by 5' ESTs or Sequences obtained Therefrom to Import Proteins Into Cells

The short core hydrophobic region (h) of signal peptides encoded by the 5'ESTS or extended cDNAs derived from SEQ ID NOs: 38-291 may also be used as a carrier to import a peptide or a protein of interest, so-called cargo, into tissue culture cells (Lin *et al.*, *J. Biol. Chem.*, **270**: 14225-14258, 1995; Du *et al.*, *J. Peptide Res.*, **51**: 235-243, 1998; Rojas *et al.*, *Nature Biotech.*, **16**: 370-375, 1998).

When cell permeable peptides of limited size (approximately up to 25 amino acids) are to be translocated across cell membrane, chemical synthesis may be used in order to add the h region to either the C-terminus or the N-terminus to the cargo peptide of interest. Alternatively, when longer peptides or proteins are to be imported into cells, nucleic acids can be genetically engineered, using techniques familiar to those skilled in the art, in order to link the extended cDNA sequence encoding the h region to the 5' or the 3' end of a DNA sequence coding for a cargo polypeptide. Such genetically engineered nucleic acids are then translated either *in vitro* or *in vivo* after transfection into appropriate cells, using conventional techniques to produce the resulting cell permeable polypeptide. Suitable hosts cells are then simply incubated with the cell permeable polypeptide which is then translocated across the membrane.

This method may be applied to study diverse intracellular functions and cellular processes. For instance, it has been used to probe functionally relevant domains of intracellular proteins and to examine protein-protein interactions involved in signal transduction pathways (Lin *et al.*, *supra*; Lin *et al.*, *J. Biol. Chem.*, **271**: 5305-5308, 1996; Rojas *et al.*, *J. Biol. Chem.*, 271: 27456-27461, 1996; Liu *et al.*, *Proc. Natl. Acad. Sci. USA*, **93**: 11819-11824, 1996; Rojas *et al.*, *Bioch. Biophys. Res. Commun.*, **234**: 675-680, 1997).

Such techniques may be used in cellular therapy to import proteins producing therapeutic effects. For instance, cells isolated from a patient may be treated with imported therapeutic proteins and then re-introduced into the host organism.

Alternatively, the h region of signal peptides of the present invention could be used in combination with a nuclear localization signal to deliver nucleic acids into cell nucleus. Such oligonucleotides may be antisense oligonucleotides or oligonucleotides designed to form

triple helixes, as described in examples 62 and 63 respectively, in order to inhibit processing and/or maturation of a target cellular RNA.

As discussed above, the cDNAs or portions thereof obtained using the 5' ESTs of the present invention can be used for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination for expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris *et al.*, *Cell* 75:791-803, 1993, the disclosure of which is hereby incorporated by reference) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins or polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins

involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation *Molecular Cloning; A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, Sambrook,. Fritsch and Maniatis eds., 1989, and *Methods in Enzymology; Guide to Molecular Cloning Techniques*, Academic Press, Berger and Kimmel eds., 1987.

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Although this invention has been described in terms of certain preferred embodiments, other embodiments which will be apparent to those of ordinary skill in the art in view of the disclosure herein are also within the scope of this invention. Accordingly, the scope of the invention is intended to be defined only by reference to the appended claims. All documents cited herein are incorporated herein by reference in their entirety.

	Search characteristic		Selection Characteristics		
Step	Program	Strand	Parameters	Identity (%)	Length (bp)
miscellanaeous	blastn	both	S=61 X=16	90	17
tRNA	fasta	both	-	80	60
rRNA	blastn	both	S=108	80	40
mtRNA	blastn	both	S=108	80	40
Procaryotic	blastn	both	S=144	90	40
Fungal	blastn	both	S=144	90	40
Alu	fasta*	both	-	70	40
L1	blastn	both	S=72	70	40
Repeats	blastn	both	S=72	70	40
Promoters	blastn	top	S=54 X=16	90	15†
Vertebrate	fasta*	both	S=108	90	30
ESTs	blastn	both	S=108 X=16	90	30
Proteins	blastx¤	top	E = 0.001	-	-

Table 1: Parameters used for each step of EST analysis

* use "Quick Fast" Database scanner
† alignement further constrained to begin closer than 10bp to EST\5' end
¤ using BLOSUM62 substitution matrix

TABLE II

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID38	new	15	Liver Fetal liver	22-6-1-A10-PU
ID39	new	13.2	Ovary Hypertrophic prostate Brain	77-16-3-B7-PU
ID40	new	13.1	Fetal brain Substantia nigra	47-47-1-F2-PU
ID41	new	11.6	Fetal kidney Cancerous prostate	58-12-2-E11-PU
ID42	new	10.7	Liver Kidney	21-4-2-D1-PU
ID43	new	9.6	Hypertrophic prostate Cancerous prostate Large intestine	77-38-4-B2-PU
ID44	new	9.4	Fetal kidney Cancerous prostate	76-10-2-B7-PU
ID45	new	9.4	Prostate Brain	33-99-2-G8-PU
ID46	new	9.1	Hypertrophic prostate Normal prostate Brain	78-32-2-C2-PU
ID47	new	9.1	Ovary Brain	26-40-3-D6-PU
ID48	new	8	Fetal kidney Brain	33-106-2-F10-PU
ID49	new	7.8	Fetal kidney Lung (cells)	58-38-1-A2-PU
ID50	new	7.4	Lymph ganglia Surrenals	62-10-3-A11-PU
ID51	new	7.4	Hypertrophic prostate Cancerous prostate	76-45-1-F5-PU
ID52	new	7.1	Fetal kidney Lung (cells) Umbilical cord Hypertrophic prostate Cancerous prostate Substantia nigra	37-10-3-D7-PU
ID53	new	6.9	Hypertrophic prostate Normal prostate Lymph ganglia Spleen	78-16-2-B12-PU
ID54	new	6.8	Fetal brain Brain	33-38-2-A4-PU
ID55	new	6.7	Heart Spleen Substantia nigra	47-25-4-A2-PU
ID56	new	6.3	Fetal brain Spleen	20-10-3-D9-PU
ID57	new	6.3	Hypertrophic prostate	84-5-1-C9-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID58	new	6.3	Thyroid Prostate Hypertrophic prostate Normal prostate	76-40-1-A8-PU
ID59	new	6.3	Cancerous prostate Fetal kidney Normal prostate Hypertrophic prostate	76-5-1-F4-PU
ID60	new	6.3	Cancerous prostate Fetal kidney Hypertrophic prostate	77-25-3-H5-PU
ID61	new	5.7	Kidney Prostate Lymph ganglia	42-1-4-H1-PU
ID62	new	5.6	Lung Brain Lymph ganglia	33-80-4-E4-PU
ID63	new	5.6	Pancreas Fetal kidney Normal prostate	58-47-2-E11-PU
ID64	new	5.6	Muscle Brain	33-56-4-F4-PU
ID65	new	5.5	Placenta Lung (cells) Colon	23-1-4-F6-PU
ID66	new	5.3	Cancerous prostate Normal prostate Cancerous prostate	76-44-2-F7-PU
ID67	new	5.2	Hypertrophic prostate	76-19-1-E9-PU
ID68	new	5.1	Cancerous prostate Colon Normal prostate	78-31-1-D12-PU
ID69	new	4.9	Kidney Prostate	20-1-4-H6-PU
ID70	new	4.9	Spleen Lymphocytes	24-3-4-C4-PU
ID71	new	4.7	Cancerous prostate Kidney	33-102-2-C9-PU
ID72	new	4.7	Brain Colon	48-47-3-A5-PU
ID73	new	4.6	Lymph ganglia Placenta	77-2-3-D1-PU
ID74	new	4.6	Hypertrophic prostate Normal prostate Thyroid Cancerous prostate	76-3-3-C7-PU
ID75	new	4.5	Substantia nigra Fetal kidney Large intestine	83-1-3-H6-PU
ID76	new	4.4	Fetal brain Brain	33-7-2-D11-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID77	new	4	Normal prostate Substantia nigra	78-28-2-G12-PU
ID78	new	3.9	Normal prostate Cancerous prostate	76-23-3-D8-PU
ID79	new	3.9	Heart Lymph ganglia	48-3-3-H9-PU
ID80	new	3.8	Brain Lung	42-2-4-B8-PU
ID81	new	3.8	Normal prostate Hypertrophic prostate	77-37-2-H1-PU
ID82	new	3.8	Lung (cells) Testis Lung	51-37-4-B1-PU
ID83	new	3.7	Ovary Lung (cells) Colon Normal prostate	23-9-4-G9-PU
ID84	new	3.5	Ovary Muscle Hypertrophic prostate	27-3-2-B6-PU
ID85	new	3.5	Normal prostate Hypertrophic prostate Cancerous prostate	76-30-3-B7-PU
ID86	ext-est-not-vrt	13.4	Ovary Prostate Cancerous prostate	76-9-4-G9-PU
ID87	ext-est-not-vrt	12.6	Normal prostate Hypertrophic prostate	78-25-4-H1-PU
ID88	ext-est-not-vrt	11.8	Fetal kidney Hypertrophic prostate	77-1-4-D10-PU
ID89	ext-est-not-vrt	11.2	Lung (cells) Normal prostate Cancerous prostate	78-37-1-A12-PU
ID90	ext-est-not-vrt	10.3	Umbilical cord Hypertrophic prostate	37-10-2-C10-PU
ID91	ext-est-not-vrt	10.1	Brain Cancerous prostate	76-16-1-H5-PU
ID92	ext-est-not-vrt	9.8	Lymphocytes Lung (cells) Umbilical cord Normal prostate	24-1-4-G11-PU
ID93	ext-est-not-vrt	9.3	Thyroid Heart Lymph ganglia Lung	48-51-2-C10-PU
ID94	ext-est-not-vrt	8.4		33-97-4-G8-PU
ID95	ext-est-not-vrt	7.8	Fetal brain Brain	33-22-1-F9-PU
ID96	ext-est-not-vrt	7.4	Ovary Liver Umbilical cord	37-7-4-E7-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Kidney Surrenals	
ID97	ext-est-not-vrt	7.2	Muscle Liver Dystrophic muscle Normal prostate Testis Cancerous prostate Lymph ganglia Large intestine	27-12-3-H8-PU
ID98	ext-est-not-vrt	7.1	Fetal kidney Ovary	58-23-4-G9-PU
ID99	ext-est-not-vrt	6.9	Placenta Fetal kidney	58-34-2-H8-PU
ID100	ext-est-not-vrt	6.7	Fetal kidney Fetal brain Umbilical cord Heart Fetal liver	37-9-1-D4-PU
ID101	ext-est-not-vrt	6.6	Fetal kidney Liver Thyroid Kidney Cancerous prostate Lung (cells) Normal prostate Lymph ganglia	58-5-3-A8-PU
ID102	ext-est-not-vrt	6.6	Cancerous prostate Normal prostate	76-35-1-A11-PU
ID103	ext-est-not-vrt	5.4	Hypertrophic prostate Lung (cells)	77-35-2-E10-PU
ID104	ext-est-not-vrt	5.4	Fetal kidney Fetal brain Normal prostate	58-52-4-D8-PU
ID105	ext-est-not-vrt	5.3	Cancerous prostate Substantia nigra	47-26-3-D2-PU
ID106	ext-est-not-vrt	5.1	Cancerous prostate Fetal brain Lung (cells) Brain	30-9-1-G8-PU
ID107	ext-est-not-vrt	4.9	Lung Brain	33-98-1-C6-PU
ID108	ext-est-not-vrt	4.5	Ovary Prostate Normal prostate Brain	78-26-1-B12-PU
ID109	ext-est-not-vrt	4.2	Fetal kidney Cancerous prostate Normal prostate	58-7-2-F8-PU
ID110	ext-est-not-vrt	3.7	Fetal kidney Ovary	58-33-1-F9-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Prostate	
			Normal prostate	
ID111	ext-est-not-vrt	3.6	Brain	33-19-1-F1-PU
			Lymph ganglia	
ID112	ext-est-not-vrt	3.5	Fetal kidney	58-14-2-D3-PU
			Liver	
			Kidney	
			Brain	
ID113	ext-est-not-vrt	3.5	Ovary	26-40-2-B2-PU
			Hypertrophic prostate	
ID114	est-not-ext	13.9	Fetal kidney	58-52-4-F10-PU
			Cancerous prostate	
			Normal prostate	
ID115	est-not-ext	13.9	Fetal kidney	58-15-1-H6-PU
			Lung (cells)	
ID116	est-not-ext	11.6	Ovary	51-29-2-B2-PU
			Dystrophic muscle	
			Cancerous prostate	
			Uterus	
			Testis	
			Lymph ganglia	
			Surrenals	
ID117	est-not-ext	11.6	Lymph ganglia	48-7-1-F2-PU
			Large intestine	
ID118	est-not-ext	11.6	Umbilical cord	37-6-1-E12-PU
			Pancreas	
ID119	est-not-ext	11.4	Heart	67-3-4-G7-PU
			Brain	
ID120	est-not-ext	11.2	Dystrophic muscle	33-35-4-F4-PU
			Brain	
ID121	est-not-ext	11	Ovary	48-14-1-A11-PU
			Heart	
			Kidney	
			Cancerous prostate	
			Lymph ganglia	
ID122	est-not-ext	10.5	Lung	37-11-1-G2-PU
			Umbilical cord	
			Normal prostate	
ID123	est-not-ext	10	Fetal kidney	58-3-4-G2-PU
			Cancerous prostate	
			Normal prostate	
			Brain	
ID124	est-not-ext	9.5	Fetal kidney	76-18-1-F6-PU
			Cancerous prostate	
			Umbilical cord	
			Normal prostate	
ID125	est-not-ext	9.5	Placenta	47-24-2-C1-PU
			Muscle	
			Substantia nigra	
ID126	est-not-ext	9.3	Ovary	37-11-4-H11-PU
			Cancerous prostate	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Umbilical cord	
			Colon	
			Normal prostate	
			Testis	
ID127	est-not-ext	9.3	Cancerous prostate	47-37-2-E3-PU
			Normal prostate	
			Substantia nigra	
ID128	est-not-ext	9.3	Spleen	27-16-1-E4-PU
			Muscle	
ID129	est-not-ext	9.3	Colon	47-5-1-G3-PU
			Substantia nigra	
ID130	est-not-ext	9.2	Ovary	57-2-4-E11-PU
			Hypertrophic prostate	
			Fetal brain	
ID131	est-not-ext	9	Cancerous prostate	76-32-1-G12-PU
			Normal prostate	
ID132	est-not-ext	8.9	Fetal kidney	77-25-1-C6-PU
			Hypertrophic prostate	
			Placenta	
			Normal prostate	
			Brain	
ID133	est-not-ext	8.8	Dystrophic muscle	37-7-2-B11-PU
			Umbilical cord	
			Brain	
ID134	est-not-ext	8.8	Fetal kidney	77-7-3-C8-PU
			Dystrophic muscle	
			Hypertrophic prostate	
			Thyroid	
			Cancerous prostate	
			Fetal brain	
			Muscle	
			Lung (cells)	
			Normal prostate	
			Brain	
			Lymph ganglia	
			Large intestine	
ID135	est-not-ext	8.7	Fetal kidney	48-7-3-G5-PU
			Prostate	
			Hypertrophic prostate	
			Spleen	
			Lung (cells)	
			Umbilical cord	
			Testis	
			Brain	
			Lymph ganglia	
ID136	est-not-ext	8.6	Fetal kidney	78-17-2-E5-PU
			Normal prostate	
ID137	est-not-ext	8.6	Placenta	33-10-4-E2-PU
			Brain	
ID138	est-not-ext	8.5	Umbilical cord	37-11-1-C7-PU
			Normal prostate	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID139	est-not-ext	8.5	Fetal kidney	26-48-1-H10-PU
			Lymphocytes	
			Ovary	
			Hypertrophic prostate	
ID140	est-not-ext	8.3	Prostate	60-13-3-F6-PU
			Cancerous prostate	
			Spleen	
			Normal prostate	
			Brain	
			Lymph ganglia	
			Large intestine	
ID141	est-not-ext	8.3	Cancerous prostate	78-22-4-A12-PU
			Normal prostate	
ID142	est-not-ext	8.1	Fetal kidney	57-28-4-B11-PU
			Ovary	
			Dystrophic muscle	
			Hypertrophic prostate	
			Cancerous prostate	
			Lung	
			Spleen	
			Placenta	
			Fetal brain	
			Normal prostate	
			Colon	
			Brain	
			Substantia nigra	
ID143	est-not-ext	8	Cancerous prostate	33-106-3-D8-PU
			Uterus	
			Lung (cells)	
			Colon	
			Brain	
			Substantia nigra	
ID144	est-not-ext	7.9	Normal prostate	23-8-3-F5-PU
			Colon	
ID145	est-not-ext	7.8	Placenta	17-1-3-H5
			Brain	
ID146	est-not-ext	7.6	Lung	33-37-2-G9-PU
			Normal prostate	
			Brain	
			Substantia nigra	
ID147	est-not-ext	7.6	Brain	51-16-4-H4-PU
			Testis	
ID148	est-not-ext	7.6	Hypertrophic prostate	33-32-3-G1-PU
			Cancerous prostate	
			Fetal brain	
			Muscle	
			Brain	
			Lymph ganglia	
			Large intestine	
			Surrenals	
ID149	est-not-ext	7.6	Fetal kidney	47-10-4-F3-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Hypertrophic prostate	
			Cancerous prostate	
			Lung (cells)	
			Umbilical cord	
			Normal prostate	
			Brain	
			Surrenals	
			Substantia nigra	
ID150	est-not-ext	7.4	Heart	51-1-3-G10-PU
			Cancerous prostate	
			Testis	
ID151	est-not-ext	7.4	Umbilical cord	33-39-4-B2-PU
			Brain	
			Lymph ganglia	
ID152	est-not-ext	7.4	Normal prostate	47-14-3-A3-PU
			Brain	
			Substantia nigra	
ID153	est-not-ext	7.4	Liver	48-53-3-H11-PU
			Lymph ganglia	
ID154	est-not-ext	7.4	Cerebellum	33-63-1-C3-PU
			Dystrophic muscle	
			Hypertrophic prostate	
			Heart	
			Uterus	
			Umbilical cord	
			Brain	
ID155	est-not-ext	7.3	Fetal kidney	53-3-4-F11-PU
			Ovary	
			Hypertrophic prostate	
			Spleen	
			Lung (cells)	
			Umbilical cord	
			Normal prostate	
			Brain	
			Substantia nigra	
ID156	est-not-ext	7.2	Fetal kidney	48-5-4-E8-PU
			Fetal brain	
			Uterus	
			Muscle	
			Umbilical cord	
			Lung (cells)	
			Colon	
			Normal prostate	
			Brain	
			Lymph ganglia	
			Fetal liver	
			Substantia nigra	
			Surrenals	
ID157	est-not-ext	7.1	Cancerous prostate	48-54-3-D2-PU
			Lymph ganglia	
			Large intestine	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID158	est-not-ext	7.1	Surrenals Prostate Hypertrophic prostate Cancerous prostate Normal prostate	78-18-3-C8-PU
ID159	est-not-ext	7.1	Normal prostate Testis	51-4-2-E10-PU
ID160	est-not-ext	7	Fetal kidney Lymphocytes Umbilical cord	24-11-1-E4-PU
ID161	est-not-ext	7	Cancerous prostate Brain	76-1-2-B8-PU
ID162	est-not-ext	6.7	Ovary Thyroid Cancerous prostate Uterus Muscle Normal prostate Testis Lymph ganglia	51-11-3-G9-PU
ID163	est-not-ext	6.7	Hypertrophic prostate Lung Brain Surrenals	77-16-4-G3-PU
ID164	est-not-ext	6.6	Fetal kidney Hypertrophic prostate	77-38-2-D5-PU
ID165	est-not-ext	6.6	Fetal kidney Cancerous prostate Brain	58-3-3-C8-PU
ID166	est-not-ext	6.5	Brain Testis	51-1-4-C1-PU
ID167	est-not-ext	6.5	Fetal kidney Brain Lymph ganglia	58-9-2-A6-PU
ID168	est-not-ext	6.3	Fetal kidney Cancerous prostate Lung (cells)	30-4-1-E7-PU
ID169	est-not-ext	6.3	Normal prostate Brain	33-51-3-H4-PU
ID170	est-not-ext	6.3	Cancerous prostate Fetal brain	57-27-3-A11-PU
ID171	est-not-ext	6.3	Hypertrophic prostate Fetal brain Normal prostate Brain	57-5-4-G1-PU
ID172	est-not-ext	6.2	Fetal kidney Normal prostate Testis	58-6-1-H4-PU
ID173	est-not-ext	6.2	Fetal kidney Liver Cancerous prostate	37-12-1-D7-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID174	est-not-ext	6.2	Umbilical cord Cancerous prostate Normal prostate Large intestine	78-13-1-H1-PU
ID175	est-not-ext	6.2	Brain Substantia nigra	33-18-3-G10-PU
ID176	est-not-ext	6.2	Normal prostate Substantia nigra	78-39-4-B9-PU
ID177	est-not-ext	6.2	Brain Substantia nigra	33-18-2-B1-PU
ID178	est-not-ext	6.1	Fetal kidney Umbilical cord Normal prostate	37-4-3-D5-PU
ID179	est-not-ext	6.1	Cerebellum Muscle Brain Substantia nigra Fetal kidney Prostate Hypertrophic prostate Cancerous prostate Lung Lung (cells) Umbilical cord Normal prostate Testis Lymph ganglia Large intestine Surrenals	58-35-3-D12-PU
ID180	est-not-ext	6.1	Fetal liver Testis	51-38-3-D10-PU
ID181	est-not-ext	6.1	Uterus Fetal liver Substantia nigra Ovary Cancerous prostate Fetal brain Normal prostate Lymph ganglia	76-14-3-G2-PU
ID182	est-not-ext	6.1	Cancerous prostate Normal prostate	76-30-1-F7-PU
ID183	est-not-ext	6	Brain Cancerous prostate	76-43-3-E11-PU
ID184	est-not-ext	6	Thyroid Pancreas Fetal kidney Normal prostate	78-41-2-H7-PU
ID185	est-not-ext	5.9	Liver Lung	59-8-1-B7-PU
ID186	est-not-ext	5.8	Brain Lung	78-37-4-E6-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID187	est-not-ext	5.8	Normal prostate Kidney Cancerous prostate	59-1-2-E4-PU
ID188	est-not-ext	5.7	Lung Umbilical cord Normal prostate	78-38-4-G2-PU
ID189	est-not-ext	5.7	Lymphocytes Spleen Uterus Substantia nigra Fetal kidney Hypertrophic prostate Cancerous prostate Normal prostate Testis	20-1-3-G5-PU
ID190	est-not-ext	5.7	Brain Fetal kidney	58-37-3-E3-PU
ID191	est-not-ext	5.7	Brain Fetal brain	33-15-1-H3-PU
ID192	est-not-ext	5.6	Lymphocytes Thyroid Spleen Uterus Substantia nigra Hypertrophic prostate Umbilical cord Normal prostate Surrenals	37-1-1-C2-PU
ID193	est-not-ext	5.6	Fetal kidney Umbilical cord Lymph ganglia	48-10-1-A8-PU
ID194	est-not-ext	5.6	Surrenals	62-1-2-D2-PU
ID195	est-not-ext	5.6	Brain Hypertrophic prostate	33-12-4-A7-PU
ID196	est-not-ext	5.6	Brain Normal prostate	78-30-4-H3-PU
ID197	est-not-ext	5.6	Cerebellum Brain Substantia nigra Fetal kidney Hypertrophic prostate Lung Fetal brain Normal prostate Lymph ganglia	47-8-4-C11-PU
ID198	est-not-ext	5.6	Thyroid Brain	84-4-2-C1-PU
ID199	est-not-ext	5.6	Brain Dystrophic muscle Lung (cells) Normal prostate	30-12-4-C2-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Testis	
ID200	est-not-ext	5.6	Placenta	1-32-0-D10
			Lung	
ID201	est-not-ext	5.5	Ovary	30-1-2-E3-PU
			Lung (cells)	
ID202	est-not-ext	5.5	Ovary	60-11-1-F1-PU
			Prostate	
			Lymph ganglia	
ID203	est-not-ext	5.5	Spleen	33-105-2-C3-PU
			Brain	
			Fetal kidney	
			Prostate	
			Hypertrophic prostate	
			Lung (cells)	
			Umbilical cord	
			Testis	
			Lymph ganglia	
ID204	est-not-ext	5.5	Cancerous prostate	76-31-4-H1-PU
			Normal prostate	
ID205	est-not-ext	5.5	Fetal kidney	30-10-3-B10-PU
			Ovary	
			Cancerous prostate	
			Umbilical cord	
			Lung (cells)	
ID206	est-not-ext	5.4	Muscle	27-3-2-E11-PU
			Fetal kidney	
			Cancerous prostate	
			Lung	
			Lymph ganglia	
ID207	est-not-ext	5.3	Placenta	31-9-2-F9-PU
			Muscle	
			Brain	
			Substantia nigra	
			Cancerous prostate	
			Umbilical cord	
ID208	est-not-ext	5.3	Brain	47-40-3-D2-PU
			Substantia nigra	
			Fetal kidney	
ID209	est-not-ext	5.3	Brain	33-77-1-F10-PU
			Substantia nigra	
			Lung	
ID210	est-not-ext	5.2	Cerebellum	51-19-3-D6-PU
			Ovary	
			Umbilical cord	
			Testis	
ID211	est-not-ext	5.2	Brain	51-6-2-F10-PU
			Hypertrophic prostate	
			Colon	
			Testis	
ID212	est-not-ext	5.2	Brain	33-72-4-C5-PU
			Fetal kidney	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Fetal brain	
			Umbilical cord	
			Normal prostate	
ID213	est-not-ext	5	Brain	33-18-3-E6-PU
			Normal prostate	
ID214	est-not-ext	5	Brain	33-5-2-E1-PU
			Substantia nigra	
			Fetal kidney	
			Umbilical cord	
			Lymph ganglia	
ID215	est-not-ext	5	Liver	76-22-3-E4-PU
			Uterus	
			Muscle	
			Heart	
			Cancerous prostate	
ID216	est-not-ext	5	Fetal kidney	51-15-2-H5-PU
			Testis	
ID217	est-not-ext	4.9	Colon	78-33-3-A9-PU
			Normal prostate	
ID218	est-not-ext	4.9	Brain	58-42-2-H11-PU
			Substantia nigra	
			Fetal kidney	
			Dystrophic muscle	
			Cancerous prostate	
			Lung	
			Lymph ganglia	
ID219	est-not-ext	4.9	Brain	33-111-3-F7-PU
			Substantia nigra	
ID220	est-not-ext	4.9	Substantia nigra	76-44-3-C5-PU
			Fetal kidney	
			Hypertrophic prostate	
			Cancerous prostate	
ID221	est-not-ext	4.9	Substantia nigra	78-40-4-B10-PU
			Normal prostate	
			Testis	
			Surrenals	
ID222	est-not-ext	4.9	Fetal kidney	78-6-3-F5-PU
			Normal prostate	
ID223	est-not-ext	4.9	Thyroid	58-48-4-E2-PU
			Brain	
			Fetal kidney	
ID224	est-not-ext	4.8	Placenta	77-38-1-F10-PU
			Hypertrophic prostate	
			Normal prostate	
ID225	est-not-ext	4.8	Lung (cells)	30-7-4-D6-PU
			Normal prostate	
ID226	est-not-ext	4.8	Cancerous prostate	48-4-2-H3-PU
			Lymph ganglia	
ID227	est-not-ext	4.8	Brain	33-77-4-E8-PU
			Dystrophic muscle	
			Normal prostate	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID228	est-not-ext	4.8	Brain Substantia nigra	33-111-2-B4-PU
ID229	est-not-ext	4.7	Normal prostate Surrenals	62-8-1-A5-PU
ID230	est-not-ext	4.7	Brain Fetal kidney	33-6-1-G11-PU
ID231	est-not-ext	4.7	Fetal liver Substantia nigra Fetal kidney Heart Cancerous prostate Umbilical cord Normal prostate	58-13-1-H2-PU
ID232	est-not-ext	4.7	Liver Brain Substantia nigra Fetal kidney Lung (cells) Testis Large intestine	58-40-2-H6-PU
ID233	est-not-ext	4.7	Brain Fetal brain	33-50-3-C3-PU
ID234	est-not-ext	4.7	Thyroid Spleen Placenta Muscle Brain Substantia nigra Fetal kidney Ovary Heart Cancerous prostate Lung Fetal brain Umbilical cord Normal prostate Colon Testis Lymph ganglia Surrenals	62-10-4-C5-PU
ID235	est-not-ext	4.6	Prostate Lung (cells)	60-16-2-F2-PU
ID236	est-not-ext	4.6	Muscle Brain Substantia nigra Fetal brain Testis	33-87-2-D2-PU
ID237	est-not-ext	4.6	Liver Brain	33-80-3-B8-PU
ID238	est-not-ext	4.5	Liver Cancerous prostate	22-12-3-D4-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID239	est-not-ext	4.5	Normal prostate Lymphocytes Spleen Uterus Placenta Muscle Brain Substantia nigra Fetal kidney Ovary Prostate Dystrophic muscle Hypertrophic prostate Heart Cancerous prostate Lung Fetal brain Lung (cells) Umbilical cord Normal prostate Colon Testis Lymph ganglia Surrenals	48-51-4-C11-PU
ID240	est-not-ext	4.5	Cerebellum Substantia nigra Normal prostate	47-15-1-H8-PU
ID241	est-not-ext	4.4	Hypertrophic prostate Lung (cells)	30-12-3-G5-PU
ID242	est-not-ext	4.4	Brain Fetal kidney Cancerous prostate Umbilical cord Normal prostate	58-4-4-D4-PU
ID243	est-not-ext	4.4	Spleen	53-3-2-D4-PU
ID244	est-not-ext	4.4	Pancreas Fetal kidney	58-54-2-H8-PU
ID245	est-not-ext	4.4	Thyroid Kidney Muscle Brain Ovary Cancerous prostate Umbilical cord Normal prostate	27-17-2-C12-PU
ID246	est-not-ext	4.4	Liver Placenta Heart Normal prostate Lymph ganglia	48-5-3-A1-PU
ID247	est-not-ext	4.4	Placenta	33-21-3-D12-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Brain	
ID248	est-not-ext	4.4	Substantia nigra	47-2-3-B3-PU
			Fetal kidney	
			Umbilical cord	
ID249	est-not-ext	4.3	Muscle	58-15-2-D7-PU
			Fetal kidney	
			Cancerous prostate	
			Lung (cells)	
ID250	est-not-ext	4.3	Substantia nigra	58-41-1-G7-PU
			Fetal kidney	
			Fetal brain	
ID251	est-not-ext	4.2	Brain	77-5-3-F3-PU
			Fetal kidney	
			Hypertrophic prostate	
			Normal prostate	
ID252	est-not-ext	4.2	Brain	33-106-2-B3-PU
			Fetal kidney	
ID253	est-not-ext	4.2		58-3-3-B2-PU
ID254	est-not-ext	4.2	Normal prostate	48-46-2-G12-PU
			Lymph ganglia	
ID255	est-not-ext	4.1	Brain	58-44-2-B3-PU
			Substantia nigra	
			Fetal kidney	
			Hypertrophic prostate	
			Lung (cells)	
			Testis	
ID256	est-not-ext	4.1	Cerebellum	47-18-4-E3-PU
			Substantia nigra	
ID257	est-not-ext	4.1	Muscle	78-21-3-F8-PU
			Substantia nigra	
			Normal prostate	
ID258	est-not-ext	4.1	Brain	33-49-1-H4-PU
			Surrenals	
ID259	est-not-ext	4.1	Brain	23-11-1-E11-PU
			Fetal kidney	
			Fetal brain	
			Normal prostate	
			Colon	
ID260	est-not-ext	4	Cerebellum	33-5-2-H4-PU
			Brain	
			Heart	
			Fetal brain	
			Normal prostate	
ID261	est-not-ext	4	Brain	78-12-4-D9-PU
			Normal prostate	
ID262	est-not-ext	4	Spleen	33-103-1-D10-PU
			Brain	
			Hypertrophic prostate	
			Normal prostate	
ID263	est-not-ext	4	Placenta	33-100-4-B7-PU
			Brain	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
			Substantia nigra	
			Hypertrophic prostate	
ID264	est-not-ext	3.9	Dystrophic muscle	29-11-2-D6-PU
			Umbilical cord	
ID265	est-not-ext	3.9	Normal prostate	78-27-3-D1-PU
ID266	est-not-ext	3.9	Brain	76-30-1-H7-PU
			Hypertrophic prostate	
			Cancerous prostate	
ID267	est-not-ext	3.9	Uterus	74-10-3-C9-PU
			Substantia nigra	
			Hypertrophic prostate	
ID268	est-not-ext	3.9	Cancerous prostate	76-19-1-A9-PU
ID269	est-not-ext	3.9	Liver	76-44-4-A6-PU
			Muscle	
			Brain	
			Cancerous prostate	
			Normal prostate	
ID270	est-not-ext	3.8	Uterus	74-2-1-H4-PU
			Brain	
			Substantia nigra	
ID271	est-not-ext	3.8	Muscle	27-21-1-H3-PU
			Lung (cells)	
ID272	est-not-ext	3.8	Placenta	33-13-3-E8-PU
			Brain	
ID273	est-not-ext	3.8	Thyroid	84-3-1-G10-PU
			Brain	
			Heart	
			Cancerous prostate	
			Fetal brain	
			Lung (cells)	
			Normal prostate	
			Testis	
			Lymph ganglia	
ID274	est-not-ext	3.7	Uterus	33-8-1-A3-PU
			Brain	
			Fetal kidney	
			Cancerous prostate	
ID275	est-not-ext	3.7	Dystrophic muscle	76-43-4-H1-PU
			Cancerous prostate	
ID276	est-not-ext	3.7	Thyroid	84-5-4-H7-PU
			Placenta	
ID277	est-not-ext	3.7	Brain	37-4-1-B2-PU
			Lung (cells)	
			Umbilical cord	
			Testis	
			Lymph ganglia	
ID278	est-not-ext	3.7	Kidney	74-11-4-A9-PU
			Placenta	
			Uterus	
			Hypertrophic prostate	
			Normal prostate	

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID279	est-not-ext	3.7	Lymph ganglia Surrenals Substantia nigra Hypertrophic prostate Cancerous prostate	77-2-2-B9-PU
ID280	est-not-ext	3.7	Fetal kidney Cancerous prostate	58-8-1-F2-PU
ID281	est-not-ext	3.7	Lymph ganglia Uterus Prostate	74-7-2-F2-PU
ID282	est-not-ext	3.6	Normal prostate Lymph ganglia Fetal kidney Umbilical cord Testis	37-2-1-H11-PU
ID283	est-not-ext	3.5	Large intestine Lymphocytes Brain Fetal kidney	58-6-1-F3-PU
ID284	est-not-ext	3.5	Normal prostate Muscle Brain Hypertrophic prostate	33-54-3-G1-PU
ID285	est-not-ext	3.5	Fetal liver Substantia nigra	47-39-2-H6-PU
ID286	est-not-ext	3.5	Brain Cancerous prostate Surrenals	76-17-1-F5-PU
ID287	est-not-ext	3.5	Placenta Muscle Heart	27-7-3-D1-PU
ID288	est-not-ext	3.5	Cancerous prostate Lung (cells) Umbilical cord Colon Liver Uterus Muscle Brain Ovary Dystrophic muscle Cancerous prostate Normal prostate Colon	74-5-1-E4-PU
ID289	est-not-ext	3.5	Large intestine Brain Cancerous prostate Fetal brain Umbilical cord Surrenals	57-20-1-F6-PU

SEQ. ID NO.	CATEGORY	VON HEIJNE SCORE	TISSUE SOURCE	INTERNAL DESIGNATION
ID290	ext-vrt-not-genomic	7.4	Spleen Hypertrophic prostate Lymph ganglia	48-25-3-A3-PU
ID291	ext-vrt-not-genomic	7	Brain Pancreas Hypertrophic prostate Normal prostate	46-1-3-F4-PU

TABLE III

SEQ. ID NO.	SIGNAL PEPTIDE
ID38	MSSWSRQRPKSPGGIQPHVSRTLFLLLLLAASAWG
ID39	MRVRIGLTLLLXAVLLSLASA
ID40	MFSHLPFDCVLLLLLLLLTRS
ID41	MGPVRLGILLFLFLAVDEAWA
ID42	MKSLSLLLAVALGLATA
ID43	MLLLLTLXLLGGPTWA
ID44	MKIGILLSLLNSVISQTLMSCNWKQQMRRMKTILIILIXIWIWCLG
ID45	MKASSGRCGLVRWLQVLLPFLLSLFPGALP
ID46	MIVDCVSSHLKKTGDGAKTFIIFLCHLLRGLHA
ID47	MAKALLFPSGRSVRVLYGAVNKERQXESVLNRACPPKANSKERRGRAVLGAELTQWSSPT TAGSCCSSCTLCARSSSXVIAPSPLVPFTSGLTSLSWLLXASCS
ID48	MAASEAAVVSSPSLKTDTSPVLETAGTVAAMAATPSARAAAAVVAAAARTGSEARVS KAALATKLLSLSGVFA
ID49	MKVGVLWLISFFTFTDG
ID50	MEFGLSWIFLAAILKGVQC
ID51	MAEPGHSHHLSARVRGRTERRIPRLWRLLLWAGTAFQ
ID52	MTADPRKGRMGLQACLLGLFALILS
ID53	MLVDGPSERPALCFLLLAVAMSFF
ID54	MAAPLVLVLVVAVTVRA
ID55	MTAAIRRQRELSILPKVTLEAMNTTVMQGFNRSERCPRDTRIVQLVFPALYTVVFLTGIL LNTLALWVFVHIPSSSTFIIYLKNTLVADLXMTLMLPFKILS
ID56	MSSVLAASHPLVLSSNAGTPGISEKDNRDPAGSSIGVLTLSHLISG
ID57	MGLAMEHGGSYARAGGSSRGCWYYLRYFFLFVSLIQFLIILGLVLFMVYG
ID58	MVEASLSVRHPEYNRPLLANDLMLIKLDESVSESDTIRSISIASQCPTAGNSCLVSGWGL LANG
ID59	MGGKQRDEDDEAYGKPVKYDPSFRGPIKNRSCTDVICCVLFLLFILG
ID60	MQKASVLLFLAWVCFLFY
ID61	MSPVLHFYVRPSGHEGAASGHTRRKLQGKLPELQGVETELCYNVNWTAEALPSAEETKKL MWLFGCPYCWMMLLGSXGSFL
ID62	MDVTPRESLSILVVAGSGGHTTEILRLLGSLSNAYS
ID63	MMGVAKLTLLRVLNLPHNSIG
ID64	MDVTPRESLSILVVAGSGGHTTEILRLLGSLSNAYS
ID65	MVLLTMIARVADG
ID66	MVPVENTEGPSLLNQKGTAVETEGXGSRHPPWARGCGMFTFLSSVXA
ID67	METFLEPNNKKLLFPVGRSWSCFA
ID68	MGFLWGLALPLFFFC
ID69	MQSTSNHLWLLSDILGQGATA
ID70	MVEICAGSVLPPYSNC
ID71	MVAPVLETSHVFCCPNRVRGVLNWXSGPRGLLAFGTSCSVVXY
ID72	MDSLRKMLISVAMLGAXAGVGYALLVIVTPGERRKQEMLKEMPLQDPRSREEAART QQLLLATLQEAATT
ID73	MRQTLPCIYFWGGLLPFGMLCASSTT
ID74	MADDLEQQSQGWLSSWLPTWRPTSMSQLKNVEARILQCLQNKFLARYVSLPNQNKI WTVTVSPEQNDRTPLVMVHGFGGGVGLWILNMDSLXARRTLHTXGLLGFGRXQG
ID75	MKVTGITILFWPLSMILLSDKIQS
ID76	MAAGRAQVPSSEQAWLEDAQVFIQKTLCPAVKEPNVQLTPLVIDCVKTVWLSQGRN QGSTLPLSYSFVSVQDLKTHQRLPCCSHLSWSSSAYQAWA
ID77	MSTCCWCTPGGAST
ID78	MPFAEDKTYKYICRNFSNFCXVDVVEILPYLPCLTA

SEQ. ID NO.	SIGNAL PEPTIDE
ID79	MAESEDRSLRIVLVGKTGSGKSATANTILGEEIFDSRIAAQAVTKNCQKASREWQGRDLL VVDTPGLFDTKESLXTTCKEIXRCIISSCPGPHAIVLVLLLGRYTEE
ID80	MAQKPLRLLACGDVEGKFDILFNRVQAIQKXSGNFDLLXCVGNFFGSTQ
ID81	MESRKDITNQEELWKMKPRRNLEEDDYLHKDTGETSMLKRPVLLHLHQTAHA
ID82	MESRKDITNQEEXWKMKPRRNLEEDDYLHKDTGETSMLKRPVLLHLHQTAHA
ID83	MAATCEISNIFSNYFSAMYSSEDSTLASVPPAATFG
ID84	MRDCPGVEXILDCSXRQKTEGCRLQAGKECVDSPVEGGQSEAPPSLVSFAVSSEGTEQ
ID85	MERQSRVMSEKDEYQFQHQGAVELLVFNFLLILTILT
ID86	MKMASSLAFLLLNFHVSLLLVQLLTPCSA
ID87	MVFLPLKWSLATMSFLLSSLLALLTVSTPSWC
ID88	MESAAALHFSRPASLLLLLLXCVHWS
ID89	MEKIPVSAFLLLVALSYTLA
ID90	MGPWGEPELLVWRPEAVASEPPVPVGLEVKLGALVLLLVLTLLCSL
ID91	MAPLLLQLAVLGAALA
ID92	MAMEGYWRFLXLLGSALLVGFLSVIFA
ID93	MAQSLALSLLILVLAFG
ID94	MEAMWLLCVALAVLA
ID95	MAPITTSREEFDEIPTVVGIFSAFGLVFTVSLFAWICC
ID96	MEGPRGWLVLCVLAISLA
ID97	MTAWEAMAPHVNPTLKDKALSPQQXXXTSPAPCXSNHHNKKHLILAFCAGVLLTLLLIAF IFL
ID98	MLCSLLLCECLLLXAGYA
ID99	MGHAMGLVXSLPVHCLTFA
ID100	MARCFSLVLLLTSIWT
ID101	MLLTRKQTCQLGILLSIHRQHSKDLQDIVATLGPRSATHPHQPAIQVLAQLAFLSQISQ
ID102	MWAFSELPMPLLINLIVSLLGFVATVTL
ID103	MFKVIQRSVGPASLSLLTFKVYA
ID104	MAKSLLKTASLSGRTKLLHQTGLSLYSTSHGFYEEEVKKTLQQFPGGSIDLQKEDNGIGI LTLNNPSRMNAFSGVMMLQLLEKVIELENWTEGKGLIVRGAKNTFSSGSDLNAVKSLGLQ RLPLISVALVQGWALG
ID105	MTSFSTSAQCSTSDSACRISPGQINXVRPKLPLLKILHAAGAQG
ID106	MDTAEEDICRVCRSEGTPEKPLYHPCVCTGSIKXVHQECLVQWLKHSRKEYCELCKHRFA FTPIYSPDMPSRLPIQDIFAGLVTSIGTAIRYWFHYTLVAFAWLGVVPLTAC
ID107	MLIMLGIFFNVHS
ID108	MGGLWRPGWRCVPFCGWRWIHPGSPTRAAERVEPFLRPEWSGTGGAERGLRWLGTWKR CSLRARHPALQPPRRPKSSNPFTRAXEEERRRXNKTTLTYVAAVAVGMLXASYA
ID109	MAAQCVTKVALNVSCANLLDKDIGSKSDPLCVLFLNTSG
ID110	MTGSNEFKLNQPPEDGISSVKFSPNTSQFLLVSSWDTSVRLYDVPANSMRLKYQHTGAVL DCAFYDPTHA
ID111	MGKHLWYPGQASAHLCWCGSHCCST
ID112	MLAVSLTVXLLGA
ID113	MSSTLAKIAEIEAEMARTQKNKATAHHLGLLKARLAKLRRELITPKGGGGGGPGEGFDWP RQVMLELDLLVFHLWG
ID114	MAAAVPKRMRGPAQAKLLPGSAIQALVGLARPLVLALLLVSAALS
ID115	MTPQSLLQTTLFLLSLLFLVQGAHG
ID116	MMVVGTGTSLALSSLLSLLLFAGMQIYSRQLASTEWLTIQGGLLGSGLFVFSLTAFNNLE NLVFGKGFQAKIFPEILLCLLLALFASG
ID117	MDWTWRVFCLLAVAPGAHS
ID118	MRIANRTRFSSPFLARGAGWTHGRGMMVVGTGTSLALXSLLSLLLFAGMQMYSRQLASTE WLTIQGGLLGSGLFVFSLTAFNNLENLVFGKGFQAKIFPEILLCLLLALFASG
ID119	MTSVSTQLSLVLMSLLLVLPVVEA

SEQ. ID NO.	SIGNAL PEPTIDE
ID120	MTPLLTLILVVLMGLPLAQA
ID121	MALLLALSLLVLWTSP
ID122	MGGLEPCSRLLLLPLLLAVSG
ID123	MEVPPPAPRSFLCRALCLFPRVFA
ID124	MDLRQFLMCLSLCTAFALS
ID125	MAGGVRPLRGLRALCRVLLFLSQFCILSGG
ID126	MAAAAWLQVLPVILLLLGAHP
ID127	MRTLFNLLWLALACSPVHT
ID128	MDVLFVAIFAVPLILG
ID129	MAAAAWLQVLPVILLLLGAHP
ID130	MRTLFNLLXLALACSPVHT
ID131	MGSKVADLLYWKDTRTSGVVFTGLMVSLLCLLHFSIVSVA
ID132	MAARWRFWCVSVTMVVALLIVCDVPSASA
ID133	MEGESTSAVLSGFVLGALA
ID134	MFAPAVMRAFRKNKTLGYGVPMLLLIVGGSFG
ID135	MAAAWXSGPSAPEAVTARLVGVLWFVSVTTGPWGAVATSAGGEESLKCEDLKVGQ YICKDPKINDATQEPVNCTNYTAHVSCFPAPNITCKDSSGNETHFTGNEVGFFKPISCRNV NGYSYKVAVALSLFLGWLGA
ID136	MRTLFNLLWLALACSPVHT
ID137	MDGQKKNWKDKVVDLLYWRDIKKTGVVFGASLFLLLSLTVFS
ID138	MVAPGLVLGLVLPLILWA
ID139	MSPSGRLCLLTIVGLILPTRG
ID140	MRIANRTRFSLPFLARGAGWTHGRGMMVVGTGTSLALSSLLSLLLFA
ID141	MVLGGCPVSYLLLCGQAALLLGNLLLLHCVSRSHS
ID142	MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVA
ID143	MVLLHVLFEHAVGYALLALKEVEEISLLQPQVEESVLNLGKFHSIVRLVAFCPFASS
ID144	MSGGRAPAVLLGGVASLLLSFVWMPALLPVASRLLLLPRVLLTMASG
ID145	MVAPVWYLVAAALLVGFILFLTRSRG
ID146	MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHG
ID147	MVGEAGRDLRRRRXXAVTAXKMAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLV RKLPPLCHG
ID148	MEALGKLKQFDAYPKTLEDFRVKTCGGATVTIVSGLLMLLLFLSELQY
ID149	MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHG
ID150	MRCLTTPMLLRALAQAARA
ID151	MRCLTTPMLLRALAQAARA
ID152	MDFITSTAILPLLFGCLGVFG
ID153	MHPAVFLSLPDLRCSLLLLVTWVFTPVTT
ID154	MASLGHILVFCVGLLTMAKA
ID155	MSGSSLPSALALSLLLVSGSLLP
ID156	MAVHDLIFWRDVKKTGFVFGTTLIMLLSLAAFSVIS
ID157	MXGSVECTXGWGHCAPSPLLLWTLLLFAAPFG
ID158	MQCFSFIKTMMILFNLLIFLCGAALLAVG
ID159	MRGSVECTWGXGHCAPSPLLLWTLLLFAAPFG
ID160	MALRLLKLAATSASA
ID161	MPSAFSVSSFPVSIPAVLTQTDWTEPWLMGLATFHALCVLLTCLSSRSYRLQIGHFLCLV ILVYC
ID162	MALPHQEPKPGDLIEIFRLGYEHWALYIXDGYVIHLAPPSEYPGAGSSSVFSVLSNSAEV KRERLEDVVGGCCYRVNNSLDHEYQPRPVEVIISSAKEMVGQKMKYSIVSRNCEHFVTQL RYGKSRCKQVEKAKVEVGVATALGILVVAGCSFA
ID163	MAASTSMVPVAVTAAVAPVLSINSDFSDLREIKKQLLLIAGLTRERGLLHSSKWSAELAF SLPALPLAEL

SEQ. ID NO.	SIGNAL PEPTIDE
ID164	MEEGGNLGGLIKMVHLLVLSGAWG
ID165	MAGPAAAFRRLGALSGAAALGFASYGAHGAXFPDAYGKELFDKANKHHFLHSLALL GVPHCRKPLWAGLLLASGTTLFCTS
ID166	MGHRFLRGLLTLLLPPPPLYT
ID167	MELLQVTILFLLPSICSSNS
ID168	MASSNTVLMRLVASAYSIA
ID169	MRSSCVLLTALVALA
ID170	MGIQTSPVLLASLGVGLVTLLGLAVG
ID171	MTLQWAAVATFLYAEIGLILIFCLPFIPPQRWQKIFSFNVWGKIATFWNKAFLTIIILLI VLFLDAVRE
ID172	MPSEGRCWETLKALRSSDKGRLCYYRDWLLRREVSGGPGGRRPFRPLATETFSLAVGTFC SREPVQSNNLHLFLDFCVYIPLSWG
ID173	MTKLAQWLWGLAILGSTWVALTTG
ID174	MLLAWVQAFLVSNMLLAEAYG
ID175	MAMHFIFSDTAVLLFHFWSVHSPAGMALSVLVLLLLAVLYE
ID176	MKQVHQCIERCHVPLAQAQALVTSELEKFQDRLARCTMHCNDKAKDSIDAGXKELQ VKQQLXVVXXSVLXTTCXS
ID177	MQMSYAIRCAFYQLLLAALMLVAMLQL
ID178	MMTQTCIILLIHTMQVCTT
ID179	MXXHLQTRPLFLTCLFWPLAAL
ID180	MAANYSSTXTRREHVKVKTSSQPGFLERLSETSGGMFVGLMAFLLSFYLIFT
ID181	MRGAHLTALEMLTAFASHIRA
ID182	MVHKPMMTQTCIILLIHTMQVCTT
ID183	MAGIKALISLSFGGAIGLMFLMLGCALP
ID184	MSLMPKMHLLFPLTLVRSFWS
ID185	MMKRAAAAAVGGALAVGAVPVVLSAMGFTGAGIAASSIAAKMMSAAAIANGGGVSA GSLVATLQSVGAAGLSTSSNILLASVGSVLG
ID186	MVTIILLLSCXFWA
ID187	MXKRAAAAAVGGALAVGAVPVVLSAMGFTGAGIAASSIAAKMMSAAAIANGGGVSA GSLVATLQSVGAAGLSTSSNILLASVGSVSG
ID188	MSQDGGXGELKHMVMSFRVSELQVLLGXXGRNKSGRKHELLAKALHLLKSSCAPSVQ MKIKELYRRRFPRKTLGPSDLSLLSLPPGTSP
ID189	MPXLLPVASRLLLLPRVLLTMASG
ID190	MVFSNNDEGLINKKLPKELLLRIFSFLDIVTLCRC
ID191	MVFSNNDEGLINKKLPKELLLRIFSFLDIVTLCRC
ID192	MASYFDEHDCEPSDPEQETRTNMLLELARSLFNRMDFEDLGLVVDWDHHLPPPAAKTVVE NLPRTVIRGSQAELKCPVCLLEFEEEETAIEMPCHHLFHSSCILPWLSKTNS
ID193	MPLILSLQVCRPATL
ID194	MLGITSCSDQQAKEGEGLEGSSTGSSSGNHGGSGGGNGHKPGCEKPGNEARGSGNLGFRT LRRLLGCLTLTLS
ID195	MARKALKLASWTSMALA
ID196	MAAAALPAWLSLQSRA
ID197	MVKIAFNTPTAVQKEEARQDVEALLSRTVRTQILTGKELRVATQEKEGSSGRCMLTLXXL SFILA
ID198	MIGSGLAGSGGAGGPSSTVTWCALXSNHVAATQASLLLSFVWMPALLP
ID199	MSGAQLXGFLFXVIVLTS
ID200	MSFFQLLMKRKELIPLVVFMTVAASGASS
ID201	MELAHSLLLNEEALA
ID202	MTSALTQGLERIPDQLGYLVLSEGAVLA
ID203	MAAAWPSGPXAPEAVTARLVGVLWFVSVTTG
ID204	MVLLTMIARVADG

SEQ. ID NO.	SIGNAL PEPTIDE
ID205	MVLLTMIARVADG
ID206	MTSQPVPNETIIVLPSNVINFSQAEKPEPTNQGQDSLKKHLHAEIKVIGTIQILCGMMVL SLGIXLASA
ID207	MASVVLALRTRTAVTSLLSPTPATA
ID208	MASVVLALRTRTAVTSLLSPTPATA
ID209	MMPSRTNLATGIPSSKVKYSRLSSTDDGYIDLQFKKTPPKIPYKAIALATVLFLIGA
ID210	MPLILSLQVCRPATL
ID211	MPLILSLQVCRPATL
ID212	MASSVGNVADSTEPTKRMLSFQGLAELAHREYQAGDFEAAERHCMQLWRQEPDNTG VLLLLSSIHFQC
ID213	MFGSAPQRPVAMTTAQRDSLLWKLAGLLREXGDVVLSGCSTLSLLTPTLQQLNHVFELHL GPWGPGQTGFVALPSHPADSPVILQLQFLFDVLQ
ID214	MSFIFEWIYNGFSSVLQFLGLYKKSGKLVFLGLDNAGKTTLLHMLKDDRLGQHVPTLHPT SEELTIAGMTLQLLILVGTSKHVAFG
ID215	MDKPCGCPPGVCDHGTGDRRDPWYSTVGLLPPVRA
ID216	MAAALKCLLTLGRWCPGLGVAPQARALAALVPGVTQ
ID217	MVARVWSLMRFLIKGSVAGGAVYLVYDQELLGPSDKSQAALQKAGEVVPPAMXQFS QYVCQQTGLQIPQLPAPPKIYFPIRDSWXAGIMTVMSALSVAPSKA
ID218	MVNELQNLXSLQGSQA
ID219	MLYMSLKYIRAFFFSIQPFLPCSS
ID220	MNLERVSNEEKLNLCRKYYLGGFAFLPFLWLVNIFWFFREAFLVPAYTEQSQIKGYVWRS AVGFLFWVIVLTSWITIFQ
ID221	MAGELQGTQAPSLRGXGLTSQDSGVNPNNSXRGREAMASGSNWLSGVNVVLVMAYG SLVFVLLFIFVKRQ
ID222	MTGFLLPPASRGTRRSCSRSRKRQTRRRRNPSSFVASCPTLLPFACVPGASPTTLA
ID223	MEEXSXPLVEFVKVLCTNQVLITARA
ID224	MVRRLXXVVAFVAPGES
ID225	MAVPGVGLLTRLNLCARRRTRVQRPIVRLLSCPGTVA
ID226	MMAAVPPGLEPWNRVRIPKAGNRSAVTVQNPGAALDLCIAAVIKECHLVILSLKSQTLDA
ID227	MASLDRVKVLVLGDSGVGKSSLVHLLCQNQVLG
ID228	MVFPAKRFCLVPSMEGVRWAFSCGTWLPSRA
ID229	MASKIGSRRWMLQLIMQLGSVLLTRC
ID230	MLSKGLKRKREEEEEKEPLAVDSWWLDPGHA
ID231	MDYSLAAALTLHGHWG
ID232	MSYITSQEMKCILHWFANWSGPQRERFLEDLVAKAVPEKLQPXLDSLEQLSVSGADDHLL SLXASYIFGISG
ID233	MPLLCQIEMEYLLLKWQMTMLQSMLCDLVSYPLLPLQQTKEANLDFPKIKVSSVTITPTR WFXLIVYLWVVSFIAS
ID234	MWFEILPGLSVMGVCLLIPGLA
ID235	MEFKLEAHRIVSISLGKIYNSRVQRGGIKLHKNLLVSLVLRXPAKS
ID236	MAVLSKEYGFVLLTGAASFIMVAHLAINVSKARKKYKVEYPIMYSTDPENGHIFNCIQRA HQNTLEVYPXFLFFLAVGGVYHPRIASGLGLXLDCWT
ID237	MDGHWSAAFSALTVTAMSSWARRRSSSSRRIPSLPGSPVCWA
ID238	MAQRLLLRRFLASVIS
ID239	MASLKPAFVNYFFLLLLEVSHLLLI
ID240	MNLERVSNEEKLNLCRKYYLGGFAFLPFLWLVNIFWFFREAFLVPAYTEQSQIKGYVWRS AVGFLFWVIVLTSWITI
ID241	MAQLGAVVAVASSFFCASLFS
ID242	MSLRNLWRDYKVLVFMVPLVGLIHL
ID243	MGWDGCKCLGVFCLLISIPTPSA

SEQ. ID NO.	SIGNAL PEPTIDE
ID244	MAASQAVEEMRTAWFWGSLGFAMSILLTFPVTIPVMMMPGTRXGFEXRXFRVDVVH MDENSLEFDMVGIDAAIANAFRRILLAEVPTMAVEKVLVYNNTSIVQDEILAHRLGLIPIHA
ID245	MAASKVKQDMPPPGGYGPIDYKRNLPRRGLSGYSMLAIGIGTLIYGHWSIMKWNRERRRL QIEDFEARIALLPLLQA
ID246	MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWWIIIDA
ID247	MMTQEPGIYTWPEKTRIICSACSSVPLPWTVLVFLTFLSIPSFV
ID248	MFLTALLWRGRIPG
ID249	MNQENPPPYPGPGPTAPYPPYPPQPMGPGXMGGPYPPPQGYPYQGYPQYGWQGGPQEPPK TTVYVVEDQRRDELGPSTCLTACWTALCCC
ID250	MASLEVSRSPRRSRRELEVRSPRQNKHSVLLPTYNEREELPLIVWLLVKSFSES
ID251	MCPTCLCAPSXXWG
ID252	MAAATGAVAASAASGQAEG
ID253	MAAMSLLXRVSVTAVAA
ID254	MAGPLQGGGARALDLLRGLPRVSLA
ID255	MATATEQWVLVEMVQALYEAPAYHLILEGILILWIIRLLFS
ID256	MEDPNPEENMXQQDSPKERSPQSPGGNICHLGAPKCTRCLITFADSKXXERHMKREHPAD FVAQKLQGVLFICFTCARS
ID257	MNVIDHVRDMAAAGLHSNVRLLSSLLLTMSNN
ID258	MQNVINTVKGKALEVAEYLTPVLKESKFKETGVITPEEFVAAGDHLVHHCPTWQWATG
ID259	MATLTFSLRKPLQRSLIRPSHLPLCCFDWRLSHYYRLPPAVRLHQQRGGRPGRSSADHWH SGVPTRILPPAHRLLCIQRLPWLLLCRG
ID260	MEKPLFPLVPLHWFGFGYTALVVSGGIVGYVKTGSVPSLAAGLLFGSXA
ID261	MASTVVAVGLTIAAAGFA
ID262	MVIRVYIASSSGSTAIKKKQQDVLGFLEANKIGFEEKDIAANEENRKWMRENVPENSRPA VQGPHAFRYKAFSFSRLLSQCRP
ID263	MSSRGHSTLPRTLMAPRMISEGDIGGIAQITSSLFLGRGSVA
ID264	MAAPGPALCLFDVDGTLT
ID265	MPLGARILFHGVFYAGGFA
ID266	MLLSIGMLMLSAT
ID267	MSLTSSSSVRVEWIAAVTIAAGTAA
ID268	MSGSNGSKENSHNKARTSPYPGSKVERSQVPNEKVGWLVEWQDYKPVEYTAVSVLA GPRWA
ID269	MAISLRSSGISVKCLSKLWMRWTVTSTTRA
ID270	MSEVRLPPLRALDDFVLGSARLGGSGS
ID271	MKLVSATAWLEECWW
ID272	MKAISVSLLRLTKLLWFFSIVLYVPLLAVCCLHS
ID273	MGSLSGLRLAAGSCFRLCERDVSXSLRLTRSSDLKRINGFCTKPQESPGAPSRTYNRVPL HKPTDWQKKILIWSGRFKKEXXIPETVSLEMLXXAKNKMRVKISYLMIALTVVGCIFM
ID274	METLYRVPFLVLECPNLKLKKPPWLHMPSAMTVYALVVVSYFLITGGIIYDVIVEPPSVG SMTDEHGHQRPVAFLAYRVNGQYIMEGLASSFLFTMGGLG
ID275	MLVLRSGLTKALA
ID276	MAAPLSVEVEFGGGAXSCLTVLRNIESLAWTGGTLG
ID277	MTHLIEYDRHRKSRLSPLQHLYLLPADHSRNAAERFPGAWFQPPTVDSEASAFVGGLPVI FWSWA
ID278	MAAAALGQIWARKLLSVPWLLC
ID279	MAVESRVTQEEIKKEPEKPIDREKTCPLLLLVFTTNNG
ID280	MRLKYQHTGAVLDCAFYDPTHA
ID281	MALLFARSLRLCRWGAKRLGVASTEAQRGVSFKLXEKTAHSSLALFRDDTGVKYGL VGLEPTKVALNVERFREWAVVLADTAVTSG
ID282	MAAAAAGTXTSQRFFQSFSDALIDEDPQAALEELTKALEQKPDDAQYYCQRAYCHILLGN YCVAVADA

SEQ. ID NO.	SIGNAL PEPTIDE
ID283	MAQLKYMENVGYAQEDRERMHRNIVSLAQNLLNFMIGSILDLWQCFLWFYIGSSLNGTRG
ID284	MSPAFRAMDVEPRAKGSFWSPLSTRSGGTHA
ID285	MADEELEALRRQRLAELQAKHGDPGDAAQQEAKHREAEMRNSILAQVLDQSARA
ID286	MSAAGARGLRATYHRLLDKVELMLPEKLRPLYNHPAGPRTVFFWAPIMKWGLVCAGL ADMARP
ID287	MSNYSVSLVGPAPWGFRLQGGKDFNMPLTISSLKDGGKAAQANVRIGDVVLSIDGINAQG MTHLEAQNKIKGCTGXLNMTLQRASA
ID288	MANPKLLGLELSEAEAIG
ID289	MIIPLLEILIIIVLNEVLLFDVNSVYKALLCTLLLHFQNI
ID290	MDIQMANNFTPPSATPQGNDCDLYAHHSTARIVMPLHYSLVFIIGLVGNLLA
ID291	MLTIVKSPQKSYLFPSSMIGIGSLPSCWA

Minimum signal peptide score	false positive rate	false negative rate	proba(0.1)	proba(0.2)
3.5	0.121	0.036	0.467	0.664
4	0.096	0.06	0.519	0.708
4.5	0.078	0.079	0.565	0.745
5	0.062	0.098	0.615	0.782
5.5	0.05	0.127	0.659	0.813
6	0.04	0.163	0.694	0.836
6.5	0.033	0.202	0.725	0.855
7	0.025	0.248	0.763	0.878
7.5	0.021	0.304	0.78	0.889
8	0.015	0.368	0.816	0.909
8.5	0.012	0.418	0.836	0.92
9	0.009	0.512	0.856	0.93
9.5	0.007	0.581	0.863	0.934
10	0.006	0.679	0.835	0.919

TABLE IV

Minimum signal peptide score	All ESTs	New ESTs	ESTs matching public EST closer than 40 bp from beginning	ESTs extending known mRNA more than 40 bp	ESTs extending public EST more than 40 bp
3.5	2674	947	599	23	150
4	2278	784	499	23	126
4.5	1943	647	425	22	112
5	1657	523	353	21	96
5.5	1417	419	307	19	80
6	1190	340	238	18	68
6.5	1035	280	186	18	60
7	893	219	161	15	48
7.5	753	173	132	12	36
8	636	133	101	11	29
8.5	543	104	83	8	26
9	456	81	63	6	24
9.5	364	57	48	6	18
10	303	47	35	6	15

TABLE V

Tissue	All ESTs	New ESTs	ESTs matching public EST closer than 40 bp from beginning	ESTs extending known mRNA more than 40 bp	ESTs extending public EST more than 40 bp
Brain	329	131	75	3	24
Cancerous prostate	134	40	37	1	6
Cerebellum	17	9	1	0	6
Colon	21	11	4	0	0
Dystrophic muscle	41	18	8	0	1
Fetal brain	70	37	16	0	1
Fetal kidney	227	116	46	1	19
Fetal liver	13	7	2	0	0
Heart	30	15	7	0	1
Hypertrophic prostate	86	23	22	2	2
Kidney	10	7	3	0	0
Large intestine	21	8	4	0	1
Liver	23	9	6	0	0
Lung	24	12	4	0	1
Lung (cells)	57	38	6	0	4
Lymph ganglia	163	60	23	2	12
Lymphocytes	23	6	4	0	2
Muscle	33	16	6	0	4
Normal prostate	181	61	45	7	11
Ovary	90	57	12	1	2
Pancreas	48	11	6	0	1
Placenta	24	5	1	0	0
Prostate	34	16	4	0	2
Spleen	56	28	10	0	1
Substantia nigra	108	47	27	1	6
Surrenals	15	3	3	1	0
Testis	131	68	25	1	8
Thyroid	17	8	2	0	2
Umbilical cord	55	17	12	1	3
Uterus	28	15	3	0	2
Non tissue-specific	568	48	177	2	28
Total	2677	947	601	23	150

TABLE VI

Description of Transcription Factor Binding Sites present on promoters isolated from SignalTag sequences

Promoter sequence P13H2 (546 bp):

Matrix	Position	Orientation	Score	Length	Sequence
CMYB_01	-502	+	0.983	9	TGTCAGTTG
MYOD_Q6	-501	-	0.961	10	CCCAACTGAC
S8_01	-444	-	0.960	11	AATAGAATTAG
S8_01	-425	+	0.966	11	AACTAAATTAG
DELTAEF1_01	-390	-	0.960	11	GCACACCTCAG
GATA_C	-364	-	0.964	11	AGATAAATCCA
CMYB_01	-349	+	0.958	9	CTTCAGTTG
GATA1_02	-343	+	0.959	14	TTGTAGATAGGACA
GATA_C	-339	+	0.953	11	AGATAGGACAT
TAL1ALPHAE47_01	-235	+	0.973	16	CATAACAGATGGTAAG
TAL1BETAE47_01	-235	+	0.983	16	CATAACAGATGGTAAG
TAL1BETAITF2_01	-235	+	0.978	16	CATAACAGATGGTAAG
MYOD_Q6	-232	-	0.954	10	ACCATCTGTT
GATA1_04	-217	-	0.953	13	TCAAGATAAAGTA
IK1_01	-126	+	0.963	13	AGTTGGGAATTCC
IK2_01	-126	+	0.985	12	AGTTGGGAATTC
CREL_01	-123	+	0.962	10	TGGGAATTCC
GATA1_02	-96	+	0.950	14	TCAGTGATATGGCA
SRY_02	-41	-	0.951	12	TAAAACAAAACA
E2F_02	-33	+	0.957	8	TTTAGCGC
MZF1_01	-5	-	0.975	8	TGAGGGGA

Promoter sequence P15B4 (861bp) :

Matrix	Position	Orientation	Score	Length	Sequence
NFY_Q6	-748	-	0.956	11	GGACCAATCAT
MZF1_01	-738	+	0.962	8	CCTGGGGA
CMYB_01	-684	+	0.994	9	TGACCGTTG
VMYB_02	-682	-	0.985	9	TCCAACGGT
STAT_01	-673	+	0.968	9	TTCCTGGAA
STAT_01	-673	-	0.951	9	TTCCAGGAA
MZF1_01	-556	-	0.956	8	TTGGGGGA
IK2_01	-451	+	0.965	12	GAATGGGATTTC
MZF1_01	-424	+	0.986	8	AGAGGGGA
SRY_02	-398	-	0.955	12	GAAAACAAAACA
MZF1_01	-216	+	0.960	8	GAAGGGGA
MYOD_Q6	-190	+	0.981	10	AGCATCTGCC
DELTAEF1_01	-176	+	0.958	11	TCCCACCTTCC
S8_01	5	-	0.992	11	GAGGCAATTAT
MZF1_01	16	-	0.988	8	AGAGGGGA

Promoter sequence P29B6 (555 bp) :

Matrix	Position	Orientation	Score	Length	Sequence
ARNT_01	-311	+	0.964	16	GGACTCACGTGCTGCT
NMYC_01	-309	+	0.965	12	ACTCACGTGCTG
USF_01	-309	+	0.985	12	ACTCACGTGCTG
USF_01	-309	-	0.985	12	CAGCACGTGAGT
NMYC_01	-309	-	0.956	12	CAGCACGTGAGT
MYCMAX_02	-309	-	0.972	12	CAGCACGTGAGT
USF_C	-307	+	0.997	8	TCACGTGC
USF_C	-307	-	0.991	8	GCACGTGA
MZF1_01	-292	-	0.968	8	CATGGGGA
ELK1_02	-105	+	0.963	14	CTCTCCGGAAGCCT
CETS1P54_01	-102	+	0.974	10	TCCGGAAGCC
AP1_Q4	-42	-	0.963	11	AGTGACTGAAC
AP1FJ_Q2	-42	-	0.961	11	AGTGACTGAAC
PADS_C	45	+	1.000	9	TGTGGTCTC

TABLE VII

CLAIMS

1. A purified or isolated nucleic acid comprising the sequence of one of SEQ ID NOs: 38-291 or comprising a sequence complementary thereto.

2. The nucleic acid of Claim 1, wherein said nucleic acid is recombinant.

3. A purified or isolated nucleic acid comprising at least 10 consecutive bases of the sequence of one of SEQ ID NOs: 38-291 or one of the sequences complementary thereto.

4. A purified or isolated nucleic acid comprising at least 15 consecutive bases of one of the sequences of SEQ ID NOs: 38-291 or one of the sequences complementary thereto.

5. The nucleic acid of Claim 4, wherein said nucleic acid is recombinant.

6. A purified or isolated nucleic acid of at least 15 bases capable of hybridizing under stringent conditions to the sequence of one of SEQ ID NOs: 38-291 or one of the sequences complementary to the sequences of SEQ ID NOs: 38-291.

7. The nucleic acid of Claim 6, wherein said nucleic acid is recombinant.

8. A purified or isolated nucleic acid encoding a human gene product, said human gene product having a sequence partially encoded by one of the sequences of SEQ ID NO: 38-291.

9. A purified or isolated nucleic acid having the sequence of one of SEQ ID NOs: 38-291 or having a sequence complementary thereto.

10. A purified or isolated nucleic acid comprising the nucleotides of one of SEQ ID NOs: 38-291 which encode a signal peptide.

11. A purified or isolated polypeptides comprising a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-291.

12. A vector encoding a fusion protein comprising a polypeptide and a signal peptide, said vector comprising a first nucleic acid encoding a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-291 operably linked to a second nucleic acid encoding a polypeptide.

13. A method of directing the extracellular secretion of a polypeptide or the insertion of a polypetide into the membrane comprising the steps of:

obtaining a vector according to Claim 12; and

introducing said vector into a host cell such that said fusion protein is secreted into the extracellular environment of said host cell or inserted into the membrane of said host cell.

14. A method of importing a polypeptide into a cell comprising contacting said cell with a fusion protein comprising a signal peptide encoded by one of the sequences of SEQ ID NOs: 38-291 operably linked to said polypeptide.

15. A method of making a cDNA encoding a human secretory protein that is partially encoded by one of SEQ ID NOs 38-291, comprising the steps of:

obtaining a cDNA comprising one of the sequences of SEQ ID NOs: 38-291;

contacting said cDNA with a detectable probe comprising at least 15 consecutive nucleotides of said sequence of SEQ ID NO: 38-291 or a sequence complementary thereto under conditions which permit said probe to hybridize to said cDNA;

identifying a cDNA which hybridizes to said detectable probe; and

isolating said cDNA which hybridizes to said probe.

16. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 15.

17. The cDNA of Claim 16 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-291.

18. A method of making a cDNA comprising one of the sequences of SEQ ID NOs: 38-291, comprising the steps of:

contacting a collection of mRNA molecules from human cells with a first primer capable of hybridizing to the polyA tail of said mRNA;

hybridizing said first primer to said polyA tail;

reverse transcribing said mRNA to make a first cDNA strand;

making a second cDNA strand complementary to said first cDNA strand using at least one primer comprising at least 15 nucleotides of one of the sequences of SEQ ID NOs 38-291; and

isolating the resulting cDNA comprising said first cDNA strand and said second cDNA strand.

19. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291 or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 18.

20. The cDNA of Claim 19 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-291.

21. The method of Claim 18, wherein the second cDNA strand is made by:

contacting said first cDNA strand with a first pair of primers, said first pair of primers comprising a second primer comprising at least 15 consecutive nucleotides of one of the sequences of SEQ ID NOs 38-291 and a third primer having a sequence therein which is included within the sequence of said first primer;

performing a first polymerase chain reaction with said first pair of nested primers to generate a first PCR product;

contacting said first PCR product with a second pair of primers, said second pair of primers comprising a fourth primer, said fourth primer comprising at least 15 consecutive nucleotides of said sequence of one of SEQ ID NO:s 38-291 , and a fifth primer, said fourth and fifth primers being capable of hybridizing to sequences within said first PCR product; and

performing a second polymerase chain reaction, thereby generating a second PCR product.

22. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein encoded by one of SEQ ID NOs 38-291, or a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 21.

23. The cDNA of Claim 22 wherein said cDNA comprises the full protein coding sequence partially included in one of the sequences of SEQ ID NOs: 38-291.

24. The method of Claim 18 wherein the second cDNA strand is made by:

contacting said first cDNA strand with a second primer comprising at least 15 consecutive nucleotides of the sequences of SEQ ID NOs: 38-291;

hybridizing said second primer to said first strand cDNA; and

extending said hybridized second primer to generate said second cDNA strand.

25. An isolated or purified cDNA encoding a human secretory protein, said human secretory protein comprising the protein partially encoded by one of SEQ ID NOs 38-291 or comprising a fragment thereof of at least 10 amino acids, said cDNA being obtainable by the method of Claim 24.

26. The cDNA of Claim 25, wherein said cDNA comprises the full protein coding sequence partially included in of one of the sequences of SEQ ID NOs: 38-291.

27. A method of making a protein comprising one of the sequences of SEQ ID NO: 292-545, comprising the steps of:

obtaining a cDNA encoding the full protein sequence partially included in one of the sequences of sequence of SEQ ID NO: 38-291;

inserting said cDNA in an expression vector such that said cDNA is operably linked to a promoter;

introducing said expression vector into a host cell whereby said host cell produces the protein encoded by said cDNA; and

isolating said protein.

28. An isolated protein obtainable by the method of Claim 27.

29. A method of obtaining a promoter DNA comprising the steps of:

obtaining DNAs located upstream of the nucleic acids of SEQ ID NO: 38-291 or the sequences complementary thereto;

screening said upstream DNAs to identify a promoter capable of directing transcription initiation; and

isolating said DNA comprising said identified promoter.

30. The method of Claim 29, wherein said obtaining step comprises chromosome walking from said nucleic acids of SEQ ID NO: 38-291 or sequences complementary thereto.

31. The method of Claim 30, wherein said screening step comprises inserting said upstream sequences into a promoter reporter vector.

32. The method of Claim 30, wherein said screening step comprises identifying motifs in said upstream DNAs which are transcription factor binding sites or transcription start sites.

33. An isolated promoter obtainable by the method of Claim 32.

34. An isolated or purified protein comprising one of the sequences of SEQ ID NO: 292-545.

35. In an array of discrete ESTs or fragments thereof of at least 15 nucleotides in length, the improvement comprising inclusion in said array of at least one of the sequences of SEQ ID NOs: 38-291, or one of the sequences complementary to the sequences of SEQ ID NOs: 38-291, or a fragment thereof of at least 15 consecutive nucleotides.

36. The array of Claim 35 including therein at least two of the sequences of SEQ ID NOs: 38-291, the sequences complementary to the sequences of SEQ ID NOs: 38-291, or fragments thereof of at least 15 consecutive nucleotides.

37. The array of Claim 35 including therein at least five of the sequences of SEQ ID NOs: 38-291, the sequences complementary to the sequences of SEQ ID NOs: 38-291, or fragments thereof of at least 15 consecutive nucleotides.

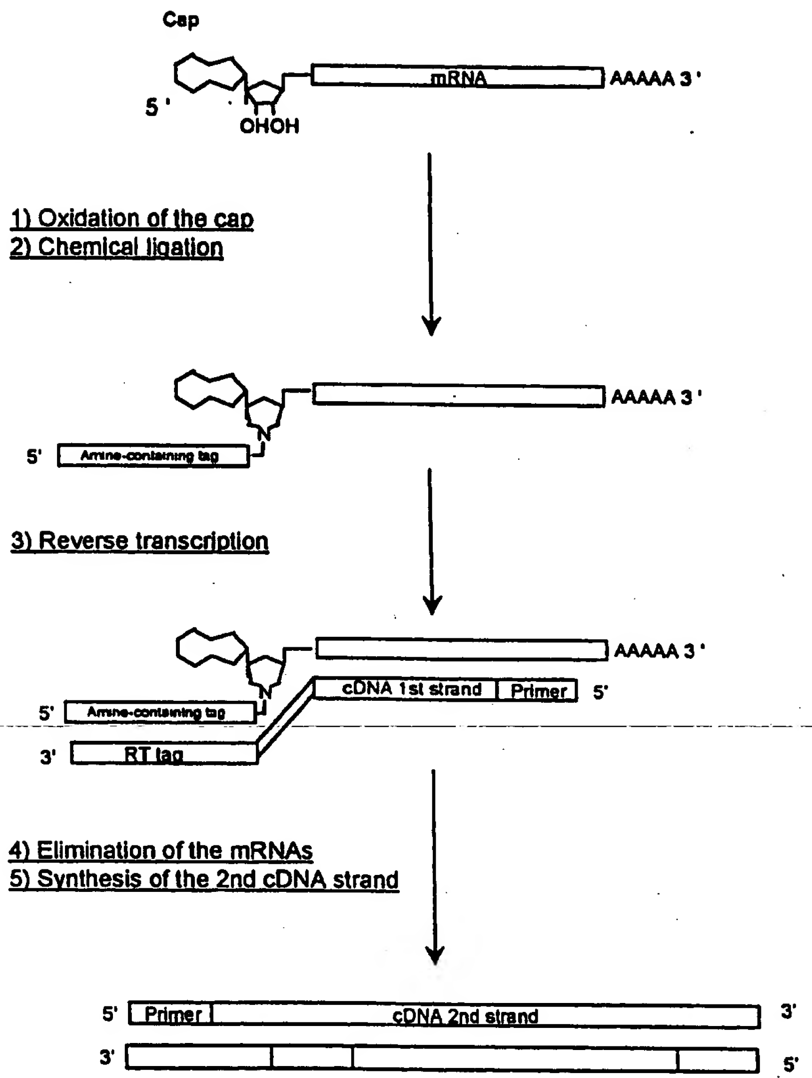

Cap
5'
OHOH
mRNA
AAAAA 3'
1) Oxidation of the cap
2) Chemical ligation
AAAAA 3'
5'
Amine-containing tag
3) Reverse transcription
AAAAA 3'
cDNA 1st strand
Primer
5'
5'
Amine-containing tag
3'
RT tag
4) Elimination of the mRNAs
5) Synthesis of the 2nd cDNA strand
5'
Primer
cDNA 2nd strand
3'
3'
5'

Figure 1

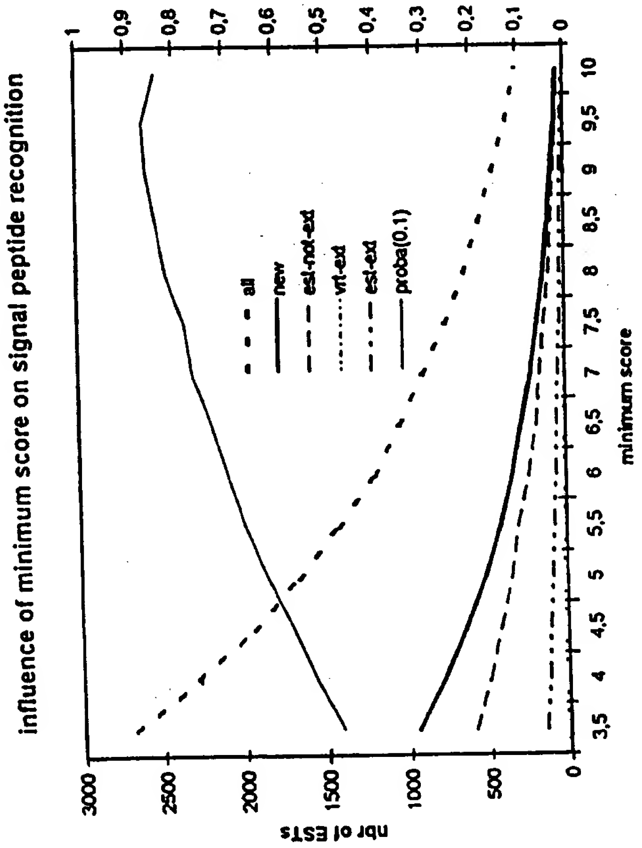

influence of minimum score on signal peptide recognition
all
new
est-not-ext
vrt-ext
est-ext
proba(0.1)
minimum score
nbr of ESTs

Figure 2

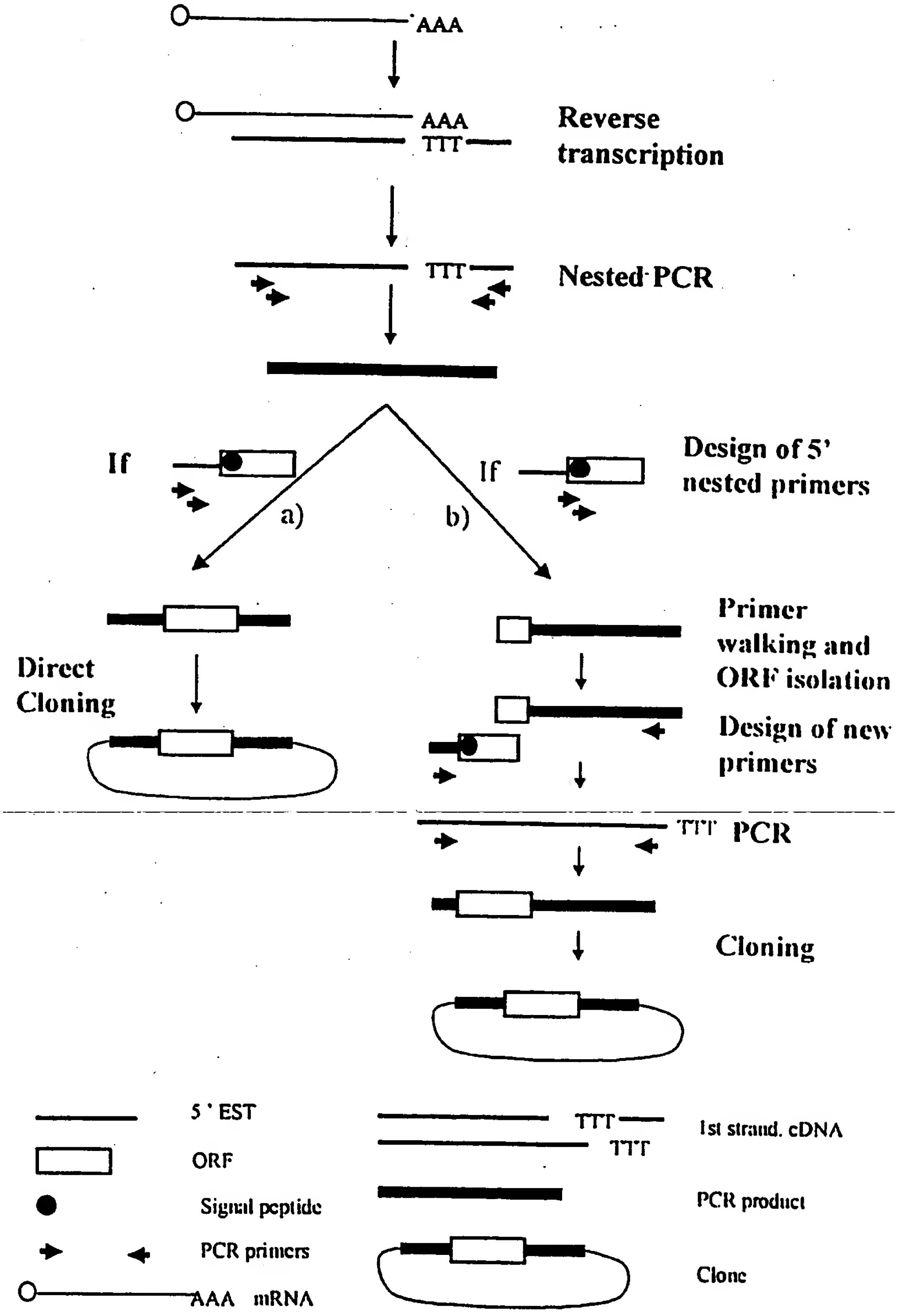

AAA
AAA
TTT
Reverse transcription
TTT
Nested PCR
If
If
Design of 5' nested primers
a)
b)
Primer walking and ORF isolation
Direct Cloning
Design of new primers
TTT
PCR
Cloning
5' EST
ORF
Signal peptide
PCR primers
AAA mRNA
TTT
TTT
1st strand. cDNA
PCR product
Clone

Figure 3

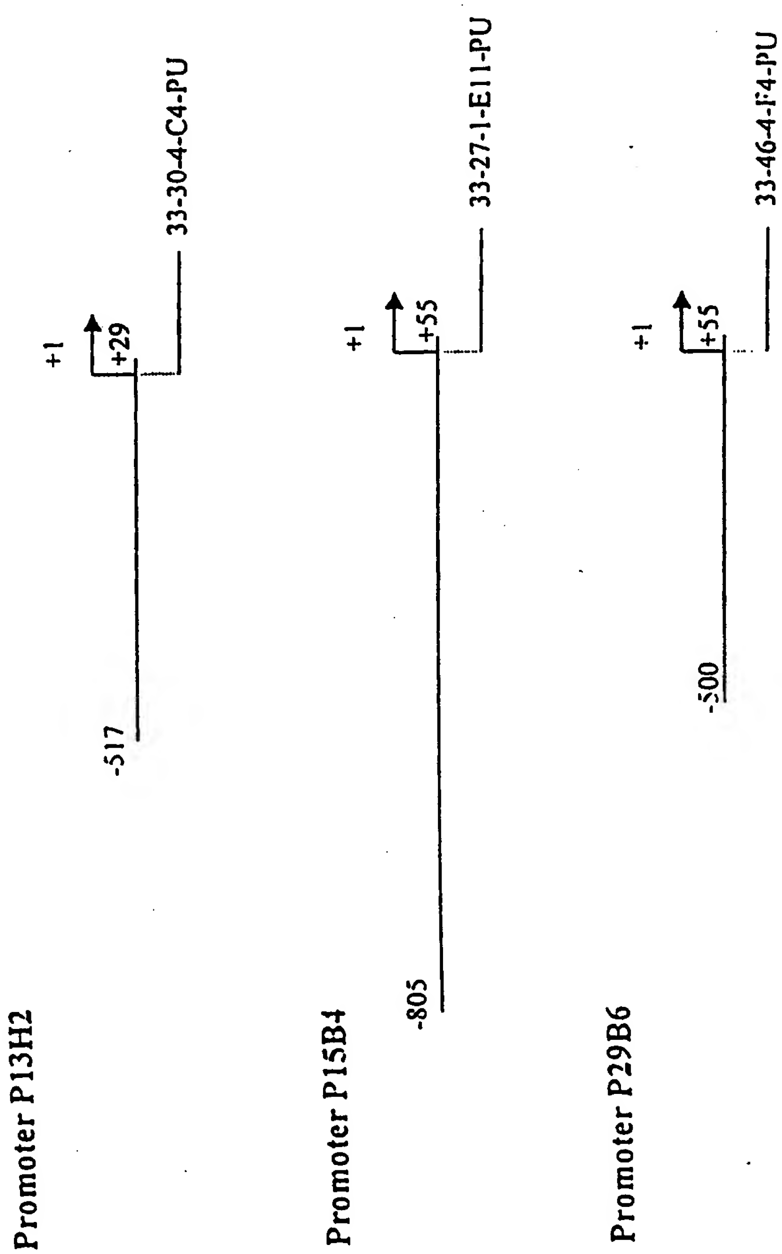

Promoter P13H2
-517
+1
+29
33-30-4-C4-PU
Promoter P15B4
-805
+1
+55
33-27-1-E11-PU
Promoter P29B6
-500
+1
+55
33-46-4-F4-PU

Figure 4

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME : GENSET SA
(B) STREET :24, RUE ROYALE
(C) CITY: PARIS
(E) COUNTRY : FRANCE
(F) POSTAL CODE (ZIP) : 75008

(ii) TITLE OF INVENTION: 5' EST FOR NON-TISSUE SPECIFIC SECRETED PROTEINS

(iii) NUMBER OF SEQUENCES: 545

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy Disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: Win95
(D) SOFTWARE: Word

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(ix) FEATURE:
(A) NAME/KEY: Cap
(B) LOCATION: 1
(D) OTHER INFORMATION: m7Gppp added to 1

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGCAUCCUAC UCCCAUCCAA UUCCACCCUA ACUCCUCCCA UCUCCAC 47

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GCAUCCUACU CCCAUCCAAU UCCACCCUAA CUCCUCCCAU CUCCAC 46

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATCAAGAATT CGCACGAGAC CATTA 25

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TAATGGTCTC GTGCGAATTC TTGAT 25

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CCGACAAGAC CAACGTCAAG GCCGC 25

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

TCACCAGCAG GCAGTGGCTT AGGAG 25

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

AGTGATTCCT GCTACTTTGG ATGGC 25

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCTTGGTCTT GTTCTGGAGT TTAGA 25

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TCCAGAATGG GAGACAAGCC AATTT 25

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

AGGGAGGAGG AAACAGCGTG AGTCC 25

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ATGGGAAAGG AAAAGACTCA TATCA 25

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

AGCAGCAACA ATCAGGACAG CACAG 25

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

ATCAAGAATT CGCACGAGAC CATTA 25

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ATCGTTGAGA CTCGTACCAG CAGAGTCACG AGAGAGACTA CACGGTACTG GTTTTTTTTT 60

TTTTTVN 67

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

CCAGCAGAGT CACGAGAGAG ACTACACGG 29

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: SINGLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: Other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CACGAGAGAG ACTACACGGT ACTGG 25

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 526 base pairs
 (B) TYPE: NUCLEIC ACID
 (C) STRANDEDNESS: DOUBLE
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: Homo Sapiens
 (F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: complement(261..376)
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 96
 region 166..281
 id N70479
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: complement(380..486)
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 97
 region 54..160
 id N70479
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: complement(110..145)
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 94
 region 403..438
 id N70479
 est

(ix) FEATURE:
 (A) NAME/KEY: other
 (B) LOCATION: complement(196..229)
 (C) IDENTIFICATION METHOD: blastn
 (D) OTHER INFORMATION: identity 94
 region 315..348
 id N70479
 est

(ix) FEATURE:
 (A) NAME/KEY: sig_peptide
 (B) LOCATION: 90..140
 (C) IDENTIFICATION METHOD: Von Heijne matrix
 (D) OTHER INFORMATION: score 8.2
 seq LLLITAILAVAVG/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

AATATRARAC AGCTACAATA TTCCAGGGCC ARTCACTTGC CATTTCTCAT AACAGCGTCA 60

```
GAGAGAAAGA ACTGACTGAR ACGTTTGAG ATG AAG AAA GTT CTC CTC CTG ATC       113
                                Met Lys Lys Val Leu Leu Leu Ile
                                        -15                 -10

ACA GCC ATC TTG GCA GTG GCT GTW GGT TTC CCA GTC TCT CAA GAC CAG       161
Thr Ala Ile Leu Ala Val Ala Val Gly Phe Pro Val Ser Gln Asp Gln
                -5                  1               5

GAA CGA GAA AAA AGA AGT ATC AGT GAC AGC GAT GAA TTA GCT TCA GGR       209
Glu Arg Glu Lys Arg Ser Ile Ser Asp Ser Asp Glu Leu Ala Ser Gly
        10                  15                  20

WTT TTT GTG TTC CCT TAC CCA TAT CCA TTT CGC CCA CTT CCA CCA ATT       257
Xaa Phe Val Phe Pro Tyr Pro Tyr Pro Phe Arg Pro Leu Pro Pro Ile
    25                  30                  35

CCA TTT CCA AGA TTT CCA TGG TTT AGA CGT AAN TTT CCT ATT CCA ATA       305
Pro Phe Pro Arg Phe Pro Trp Phe Arg Arg Xaa Phe Pro Ile Pro Ile
40                  45                  50                  55

CCT GAA TCT GCC CCT ACA ACT CCC CTT CCT AGC GAA AAG TAAACAARAA        354
Pro Glu Ser Ala Pro Thr Thr Pro Leu Pro Ser Glu Lys
                60                  65

GGAAAAGTCA CRATAAACCT GGTCACCTGA AATTGAAATT GAGCCACTTC CTTGAARAAT     414

CAAAATTCCT GTTAATAAAA RAAAAACAAA TGTAATTGAA ATAGCACACA GCATTCTCTA     474

GTCAATATCT TTAGTGATCT TCTTTAATAA ACATGAAAGC AAAAAAAAAA AA             526
```

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..17
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.2
seq LLLITAILAVAVG/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

```
Met Lys Lys Val Leu Leu Leu Ile Thr Ala Ile Leu Ala Val Ala Val
 1               5                  10                  15

Gly
```

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 822 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 260..464
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 153..357
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 118..184
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 98..164
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 56..113
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 35..92
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 454..485
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 348..379
id H57434
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 118..545
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..428
id N27248
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 65..369
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 41..345
id H94779
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 61..399
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 6..344
id H09880
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 408..458
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 355..405
id H09880
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 60..399
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 56..395
id H29351
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 393..432
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 391..430
id H29351
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 346..408
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq SFLPSALVIWTSA/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ACTCCTTTTA GCATAGGGGC TTCGGCGCCA GCGGCCAGCG CTAGTCGGTC TGGTAAGTGC 60

CTGATGCCGA GTTCCGTCTC TCGCGTCTTT TCCTGGTCCC AGGCAAAGCG GASGNAGATC 120

CTCAAACGGC CTAGTGCTTC GCGCTTCCGG AGAAAATCAG CGGTCTAATT AATTCCTCTG 180

GTTTGTTGAA GCAGTTACCA AGAATCTTCA ACCCTTTCCC ACAAAAGCTA ATTGAGTACA 240

```
CGTTCCTGTT GAGTACACGT TCCTGTTGAT TTACAAAAGG TGCAGGTATG AGCAGGTCTG      300

AAGACTAACA TTTTGTGAAG TTGTAAAACA GAAAACCTGT TAGAA ATG TGG TGG TTT      357
                                                  Met Trp Trp Phe
                                                      -20

CAG CAA GGC CTC AGT TTC CTT CCT TCA GCC CTT GTA ATT TGG ACA TCT        405
Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val Ile Trp Thr Ser
        -15                 -10                 -5

GCT GCT TTC ATA TTT TCA TAC ATT ACT GCA GTA ACA CTC CAC CAT ATA        453
Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala Val Thr Leu His His Ile
    1               5                   10                  15

GAC CCG GCT TTA CCT TAT ATC AGT GAC ACT GGT ACA GTA GCT CCA RAA        501
Asp Pro Ala Leu Pro Tyr Ile Ser Asp Thr Gly Thr Val Ala Pro Xaa
                20                  25                  30

AAA TGC TTA TTT GGG GCA ATG CTA AAT ATT GCG GCA GTT TTA TGT CAA        549
Lys Cys Leu Phe Gly Ala Met Leu Asn Ile Ala Ala Val Leu Cys Gln
            35                  40                  45

AAA TAGAAATCAG GAARATAATT CAACTTAAAG AAKTTCATTT CATGACCAAA             602
Lys

CTCTTCARAA ACATGTCTTT ACAAGCATAT CTCTTGTATT GCTTTCTACA CTGTTGAATT      662

GTCTGGCAAT ATTTCTGCAG TGGAAAATTT GATTTARMTA GTTCTTGACT GATAAATATG      722

GTAAGGTGGG CTTTTCCCCC TGTGTAATTG GCTACTATGT CTTACTGAGC CAAGTTGTAW      782

TTTGAAATAA AATGATATGA GAGTGACACA AAAAAAAAAA                            822
```

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..21
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq SFLPSALVIWTSA/AF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu Val
1               5                   10                  15
```

```
Ile Trp Thr Ser Ala
            20


(2) INFORMATION FOR SEQ ID NO: 21:

      (i) SEQUENCE CHARACTERISTICS:
            (A) LENGTH: 405 base pairs
            (B) TYPE: NUCLEIC ACID
            (C) STRANDEDNESS: DOUBLE
            (D) TOPOLOGY: LINEAR

      (ii) MOLECULE TYPE: CDNA

      (vi) ORIGINAL SOURCE:
            (A) ORGANISM: Homo Sapiens
            (F) TISSUE TYPE: Testis

      (ix) FEATURE:
            (A) NAME/KEY: other
            (B) LOCATION: complement(103..398)
            (C) IDENTIFICATION METHOD: blastn
            (D) OTHER INFORMATION:  identity 96
                                    region 1..296
                                    id AA442893
                                    est

      (ix) FEATURE:
            (A) NAME/KEY: sig_peptide
            (B) LOCATION: 185..295
            (C) IDENTIFICATION METHOD: Von Heijne matrix
            (D) OTHER INFORMATION:  score 5.9
                                    seq LSYASSALSPCLT/AP

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ATCACCTTCT TCTCCATCCT TSTCTGGGCC AGTCCCCARC CCAGTCCCTC TCCTGACCTG     60

CCCAGCCCAA GTCAGCCTTC AGCACGCGCT TTTCTGCACA CAGATATTCC AGGCCTACCT    120

GGCATTCCAG GACCTCCGMA ATGATGCTCC AGTCCCTTAC AAGCGCTTCC TGGATGAGGG    180

TGGC ATG GTG CTG ACC ACC CTC CCC TTG CCC TCT GCC AAC AGC CCT GTG     229
     Met Val Leu Thr Thr Leu Pro Leu Pro Ser Ala Asn Ser Pro Val
             -35                 -30                 -25

AAC ATG CCC ACC ACT GGC CCC AAC AGC CTG AGT TAT GCT AGC TCT GCC      277
Asn Met Pro Thr Thr Gly Pro Asn Ser Leu Ser Tyr Ala Ser Ser Ala
        -20                 -15                 -10

CTG TCC CCC TGT CTG ACC GCT CCA AAK TCC CCC CGG CTT GCT ATG ATG      325
Leu Ser Pro Cys Leu Thr Ala Pro Xaa Ser Pro Arg Leu Ala Met Met
    -5                   1                   5                  10

CCT GAC AAC TAAATATCCT TATCCAAATC AATAAARWRA RAATCCTCCC TCCARAAGGG   384
Pro Asp Asn
```

TTTCTAAAAA CAAAAAAAAA A 405

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..37
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.9
seq LSYASSALSPCLT/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

```
Met Val Leu Thr Thr Leu Pro Leu Pro Ser Ala Asn Ser Pro Val Asn
  1               5                  10                  15

Met Pro Thr Thr Gly Pro Asn Ser Leu Ser Tyr Ala Ser Ser Ala Leu
            20                  25                  30

Ser Pro Cys Leu Thr
        35
```

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 496 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 149..331
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..183
id AA397994
est

(ix) FEATURE:

```
        (A) NAME/KEY: other
        (B) LOCATION: 328..485
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 179..336
                                id AA397994
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(182..496)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 14..328
                                id AA399680
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 196..240
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 5.5
                                seq ILSTVTALTFAXA/LD

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

AAAAAATTGG TCCCAGTTTT CACCCTGCCG CAGGGCTGGC TGGGGAGGGC AGCGGTTTAG     60

ATTAGCCGTG GCCTAGGCCG TTTAACGGGG TGACACGAGC NTGCAGGGCC GAGTCCAAGG    120

CCCGGAGATA GGACCAACCG TCAGGAATGC GAGGAATGTT TTTCTTCGGA CTCTATCGAG    180

GCACACAGAC AGACC ATG GGG ATT CTG TCT ACA GTG ACA GCC TTA ACA TTT     231
                 Met Gly Ile Leu Ser Thr Val Thr Ala Leu Thr Phe
                 -15                 -10                 -5

GCC ARA GCC CTG GAC GGC TGC AGA AAT GGC ATT GCC CAC CCT GCA AGT      279
Ala Xaa Ala Leu Asp Gly Cys Arg Asn Gly Ile Ala His Pro Ala Ser
            1               5                   10

GAG AAG CAC AGA CTC GAG AAA TGT AGG GAA CTC GAG ASC ASC CAC TCG      327
Glu Lys His Arg Leu Glu Lys Cys Arg Glu Leu Glu Xaa Xaa His Ser
    15                  20                  25

GCC CCA GGA TCA ACC CAS CAC CGA AGA AAA ACA ACC AGA AGA AAT TAT      375
Ala Pro Gly Ser Thr Xaa His Arg Arg Lys Thr Thr Arg Arg Asn Tyr
30                  35                  40                  45

TCT TCA GCC TGAAATGAAK CCGGGATCAA ATGGTTGCTG ATCARAGCCC ATATTTAAAT   434
Ser Ser Ala

TGGAAAAGTC AAATTGASCA TTATTAAATA AAGCTTGTTT AATATGTCTC AAACAAAAAA    494

AA                                                                   496


(2) INFORMATION FOR SEQ ID NO: 24:
```

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..15
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq ILSTVTALTFAXA/LD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

```
Met Gly Ile Leu Ser Thr Val Thr Ala Leu Thr Phe Ala Xaa Ala
  1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 623 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 49..96
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.1
seq LVLTLCTLPLAVA/SA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

```
AAAGATCCCT GCAGCCCGGC AGGAGAGAAG GCTGAGCCTT CTGGCGTC ATG GAG AGG       57
                                                     Met Glu Arg
                                                         -15

CTC GTC CTA ACC CTG TGC ACC CTC CCG CTG GCT GTG GCG TCT GCT GGC       105
Leu Val Leu Thr Leu Cys Thr Leu Pro Leu Ala Val Ala Ser Ala Gly
            -10                 -5                  1

TGC GCC ACG ACG CCA GCT CGC AAC CTG AGC TGC TAC CAG TGC TTC AAG       153
Cys Ala Thr Thr Pro Ala Arg Asn Leu Ser Cys Tyr Gln Cys Phe Lys
    5                   10                  15
```

```
GTC AGC AGC TGG ACG GAG TGC CCG CCC ACC TGG TGC AGC CCG CTG GAC        201
Val Ser Ser Trp Thr Glu Cys Pro Pro Thr Trp Cys Ser Pro Leu Asp
20                  25                  30                  35

CAA GTC TGC ATC TCC AAC GAG GTG GTC GTC TCT TTT AAA TGG AGT GTA        249
Gln Val Cys Ile Ser Asn Glu Val Val Val Ser Phe Lys Trp Ser Val
                40                  45                  50

CGC GTC CTG CTC AGC AAA CGC TGT GCT CCC AGA TGT CCC AAC GAC AAC        297
Arg Val Leu Leu Ser Lys Arg Cys Ala Pro Arg Cys Pro Asn Asp Asn
            55                  60                  65

ATG AAK TTC GAA TGG TCG CCG GCC CCC ATG GTG CAA GGC GTG ATC ACC        345
Met Xaa Phe Glu Trp Ser Pro Ala Pro Met Val Gln Gly Val Ile Thr
        70                  75                  80

AGG CGC TGC TGT TCC TGG GCT CTC TGC AAC AGG GCA CTG ACC CCA CAG        393
Arg Arg Cys Cys Ser Trp Ala Leu Cys Asn Arg Ala Leu Thr Pro Gln
    85                  90                  95

GAG GGG CGC TGG GCC CTG CRA GGG GGG CTC CTG CTC CAG GAC CCT TCG        441
Glu Gly Arg Trp Ala Leu Xaa Gly Gly Leu Leu Leu Gln Asp Pro Ser
100                 105                 110                 115

AGG GGC ARA AAA ACC TGG GTG CGG CCA CAG CTG GGG CTC CCA CTC TGC        489
Arg Gly Xaa Lys Thr Trp Val Arg Pro Gln Leu Gly Leu Pro Leu Cys
                120                 125                 130

CTT CCC AWT TCC AAC CCC CTC TGC CCA RGG GAA ACC CAG GAA GGA            534
Leu Pro Xaa Ser Asn Pro Leu Cys Pro Xaa Glu Thr Gln Glu Gly
            135                 140                 145

TAACACTGTG GGTGCCCCCA CCTGTGCATT GGGACCACRA CTTCACCCTC TTGGARACAA      594

TAAACTCTCA TGCCCCCAAA AAAAAAAAA                                        623
```

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 1..16
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.1
seq LVLTLCTLPLAVA/SA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Met Glu Arg Leu Val Leu Thr Leu Cys Thr Leu Pro Leu Ala Val Ala

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 318 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 64..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..219
id R93883
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 281..320
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 219..258
id R93883
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 103..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 31..210
id R84338
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 281..320
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 210..249
id R84338
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 72..108
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..37
id R84338
est

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 17..259
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..243
id HUM404F03B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 20..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..263
id W05476
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 21..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..262
id R33542
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 12..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 8..278
id T85491
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 151..222
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.4
seq LMSLLLVLPVVEA/VE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 119:

```
ADTCCTGTAA TGGCTGCTTC CTAGAAGGTC GTGTCACGTG GAACCTCTTA ATCTCAGCAT    60

CCGGAGCTCC AGGAAGGGAA AATTTCAAGT CAGATAGAAT TCTATATATA CCATTTCTTT   120

GGAACCTTCA GCCCTCAAGA TTCCAACATC ATG ACC TCA GTT TCA ACA CAG TTG    174
                                 Met Thr Ser Val Ser Thr Gln Leu
                                                 -20

TCC TTA GTC CTC ATG TCA CTG CTT TTG GTG CTG CCT GTT GTG GAA GCA     222
Ser Leu Val Leu Met Ser Leu Leu Leu Val Leu Pro Val Val Glu Ala
    -15                 -10                 -5

GTA GAA GCC GGT GAT GCA ATC GCC CTT TTG TTA GGT GTG GTT CTC AGC     270
```

```
Val Glu Ala Gly Asp Ala Ile Ala Leu Leu Leu Gly Val Val Leu Ser
 1               5                  10                  15

ATT ACA GGC ATT GTG CCT GCT TGG GGG TAT ATG CAY GGG           309
Ile Thr Gly Ile Val Pro Ala Trp Gly Tyr Met His Gly
             20                  25
```

(2) INFORMATION FOR SEQ ID NO: 120:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 361 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 95..363
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 60..328
id H19572
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 140..290
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 106..256
id H46195
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 95..148
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 62..115
id H46195
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(207..316)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 183..292
id H46196
est

(ix) FEATURE:
(A) NAME/KEY: other

```
        (B) LOCATION: complement(314..363)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 137..186
                                id H46196
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(172..212)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 95
                                region 288..328
                                id H46196
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(237..287)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 92
                                region 239..289
                                id H19490
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(284..317)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 208..241
                                id H19490
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(331..363)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 93
                                region 160..192
                                id H19490
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 263..322
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 11.2
                                seq ILVVLMGLPLAQA/LD

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 120:

AAGACACGCC TACGATTAGA CTCAGGCAGG CACCTACCGG CGAGCGGCCG CRVGTGACTC    60

CCAGGCGCGG CGGTACCTCA CGGTGGTGAA GGTCACAGGG TTGCAGCACT CCCAGTAGAC   120

CAGGAGCTCC GGGAGGCAGG GCCGGCCCCA CGTCCTCTGC GCACCACCCT GAGTTGGATC   180

CTCTGTGCGC CACCCCTGAG TTGGATCCAG GGCTAGCTGC TGTTGACCTC CCCACTCCCA   240
```

```
CGCTGCCCTC CTGCCTGCAG CC ATG ACG CCC CTG CTC ACC CTG ATC CTG GTG    292
                         Met Thr Pro Leu Leu Thr Leu Ile Leu Val
                         -20                 -15

GTC CTC ATG GGC TTA CCT CTG GCC CAG GCC TTG GAC TGC CAC GTG TGT    340
Val Leu Met Gly Leu Pro Leu Ala Gln Ala Leu Asp Cys His Val Cys
-10                 -5                  1                   5

NCC TAC AAC GGA GAC AAC TGC                                        361
Xaa Tyr Asn Gly Asp Asn Cys
            10
```

(2) INFORMATION FOR SEQ ID NO: 121:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 510 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 20..372
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..353
id W05519
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 368..423
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 348..403
id W05519
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 17..260
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 21..264
id T97490
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 231..341
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 287..347

id T97490
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 16..315
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..300
id HUML12811
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 16..275
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..260
id HUML13801
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 139..186
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11
seq LLALSLLVLWTSP/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 121:

```
AATTCCCAGC CTCACATCAC TCACACCTTG CATTTCACCC CTGCATCCCA GTCGCCCTGC    60

AGCCTCACAC AGATCCTGCA CACACCCAGA CAGCTGGCGC TCACACATTC ACCGTTGGCC   120

TGCCTCTGTT CACCCTCC ATG GCC CTG CTA CTG GCC CTC AGC CTG CTA GTT     171
                    Met Ala Leu Leu Leu Ala Leu Ser Leu Leu Val
                        -15                 -10

CTC TGG ACT TCC CCA GCC CCA ACT CTG AGT GGC ACC AAT GAT GCT GAA     219
Leu Trp Thr Ser Pro Ala Pro Thr Leu Ser Gly Thr Asn Asp Ala Glu
 -5                  1               5                  10

GAC TGC TGC CTG TCT GTG ACC CAG AAA CCC ATC CCT GGG TAC ATC GTG     267
Asp Cys Cys Leu Ser Val Thr Gln Lys Pro Ile Pro Gly Tyr Ile Val
            15                  20                  25

AGG AAC TTC CAC TAC CTT CTC ATC AAG GAT GGC TGC AGG GTG CCT GCT     315
Arg Asn Phe His Tyr Leu Leu Ile Lys Asp Gly Cys Arg Val Pro Ala
        30                  35                  40

GTA GTG TTC ACC ACA CTG AGG GGC CGC CAG CTC TGT GCA CCC CCA GAC     363
Val Val Phe Thr Thr Leu Arg Gly Arg Gln Leu Cys Ala Pro Pro Asp
    45                  50                  55

CAG CCC TGG GTA GAA CGC ATC ATC CAG AGA CTG CAG AGG ACC TCA GCC     411
Gln Pro Trp Val Glu Arg Ile Ile Gln Arg Leu Gln Arg Thr Ser Ala
 60                  65                  70                  75

AAG ATG AAR MGC CGM AGC AGT KAA CCT ATG AMC GTG MAG AGG GAR CCG     459
Lys Met Lys Xaa Arg Ser Ser Xaa Pro Met Xaa Val Xaa Arg Glu Pro
```

```
                    80                      85                      90

GAG TCC GAG TCA AGC ATT GTG AAT KAT TAC CTA MCT GGG GAA CGA RGA      507
Glu Ser Glu Ser Ser Ile Val Asn Xaa Tyr Leu Xaa Gly Glu Arg Xaa
             95                     100                     105

AGG                                                                  510
Arg
```

(2) INFORMATION FOR SEQ ID NO: 122:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 382 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 152..287
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 91..226
id W60940
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 108..160
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 48..100
id W60940
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 60..106
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..47
id W60940
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 152..316
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 90..254
id H39980
est

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 62..160
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 1..99
                             id H39980
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 308..384
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 94
                             region 247..323
                             id H39980
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(148..292)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 282..426
                             id N41026
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(283..384)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 93
                             region 191..292
                             id N41026
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 66..160
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                             region 8..102
                             id R49793
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 199..271
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                             region 141..213
                             id R49793
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 152..199
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 91
                             region 93..140
                             id R49793
```

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 18..160
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..143
id W74783
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 190..253
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 173..236
id W74783
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 74..136
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.5
seq RLLLLPLLLAVSG/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 122:

```
AATTTCACTT GCCTGGACGC TGCGCCACAT CCCACCGGCC CTTACACTGT GGTGTCCAGC    60

AGCATCCGGC TTC ATG GGG GGA CTT GAA CCC TGC AGC AGG CTC CTG CTC      109
               Met Gly Gly Leu Glu Pro Cys Ser Arg Leu Leu Leu
                   -20                 -15                 -10

CTG CCT CTC CTG CTG GCT GTA AGT GGT CTC CGT CCT GTC CAG GCC CAG     157
Leu Pro Leu Leu Leu Ala Val Ser Gly Leu Arg Pro Val Gln Ala Gln
                -5                   1               5

GCC CAG AGC GAT TGC AGT TGC TCT ACG GTG AGC CCG GGC GTG CTG GCA     205
Ala Gln Ser Asp Cys Ser Cys Ser Thr Val Ser Pro Gly Val Leu Ala
         10                  15                  20

GGG ATC GTG ATG GGA GAC CTG GTG CTG ACA GTG CTC ATT GCC CTG GCC     253
Gly Ile Val Met Gly Asp Leu Val Leu Thr Val Leu Ile Ala Leu Ala
     25                  30                  35

GTG TAC TTC CTG GGC CGG CTG GTC CCT CGG GGG CGA GGG GCT GCG GAG     301
Val Tyr Phe Leu Gly Arg Leu Val Pro Arg Gly Arg Gly Ala Ala Glu
 40                  45                  50                  55

GCA SNG ACC CGG AAA CAG CGT ATC ACT GAG ACC GGG TCG CCT TAT CAG     349
Ala Xaa Thr Arg Lys Gln Arg Ile Thr Glu Thr Gly Ser Pro Tyr Gln
                 60                  65                  70

GAG CTC CAG GGT CAG AGG TCG GAT GTC TAC AGC                         382
Glu Leu Gln Gly Gln Arg Ser Asp Val Tyr Ser
             75                  80
```

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 423 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..196
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 13..155
id N41450
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 193..332
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 153..292
id N41450
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 327..425
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 288..386
id N41450
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 204..332
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 202..330
id W76359
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..124
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 54..124
id W76359
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..53
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 3..54
id W76359
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 327..370
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 326..369
id W76359
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 164..196
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 162..194
id W76359
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 133..163
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 132..162
id W76359
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..128
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 46..120
id W04321
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 9..54
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 2..47
id W04321
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 164..201
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 153..190
id W04321

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 30..94
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..65
id H75454
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 230..307
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq VLCTNQVLITARA/VP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

```
AACTTCCAAG TTGTAGTGTT GTTGTTTTCA GCCTGCTGCT GCTGCTGCTA TTGCGGCTAG    60

GGGAACCGTC GTGGGGAAGG ATGGTGTGCG AAAAATGTGA AAAGAAACTT GGTACTGTTA   120

TCACTCCAGA TACATGGAAA GATGGTGCTA GGAATACCAC AGAAAGTGGT GGAAGAAAGC   180

TGAATGAAAA TAAAGCTTTG RCTTCAAAAA AAGCCAGAAT TGAWCCATA ATG GAA GAA   238
                                                      Met Glu Glu
                                                          -25

WTA AGT KCT CCA CTT GTA GAA TTT GTA AAA GTT TTG TGC ACC AAC CAG       286
Xaa Ser Xaa Pro Leu Val Glu Phe Val Lys Val Leu Cys Thr Asn Gln
            -20                 -15                 -10

GTT CTC ATT ACT GCC AGG GCT GTG CCT ACA AAA AAG GCA TCT GTG CGA       334
Val Leu Ile Thr Ala Arg Ala Val Pro Thr Lys Lys Ala Ser Val Arg
         -5                   1               5

TGT GTG GMA AAA AGG TTT TGG ATA CCA AAA ACT ACA AGC AAA CAT CTG       382
Cys Val Xaa Lys Arg Phe Trp Ile Pro Lys Thr Thr Ser Lys His Leu
 10                  15                  20                  25

TCT AGA TGT ATT GAT GGA ATT TCT GGC TTT CTA AAT GAT TTT ACT TTC       430
Ser Arg Cys Ile Asp Gly Ile Ser Gly Phe Leu Asn Asp Phe Thr Phe
                 30                  35                  40

TGC CTT GAA TTT TCA AGG CAT AGA TGT                                   457
Cys Leu Glu Phe Ser Arg His Arg Cys
            45                  50
```

(2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 372 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

```
(vi) ORIGINAL SOURCE:
     (A) ORGANISM: Homo Sapiens
     (F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 125..367
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 119..361
                             id AA242967
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 6..125
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                             region 1..120
                             id AA242967
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 125..261
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                             region 124..260
                             id C18969
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 2..125
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 2..125
                             id C18969
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 253..311
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                             region 251..309
                             id C18969
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 125..367
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 101..343
                             id N40141
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
```

```
        (B) LOCATION: 24..125
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 1..102
                                id N40141
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 125..329
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 122..326
                                id R78319
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 9..125
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 7..123
                                id R78319
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(125..367)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 112..354
                                id N27018
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(73..125)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 94
                                region 353..405
                                id N27018
                                est

   (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 106..156
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.8
                                seq LXXVVAFVAPGES/QQ

   (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:


ATTCTTTCTT CGCCAGGCTC TCTGCTGACT CAAGTTCTTC AGTTCACGAT CTTCTAGTTG   60

CAGCGATGAG TGCACGAGTG AGATCAAGAT CCAGAGGAAG AGGAG ATG GTC AGG AGG  117
                                                  Met Val Arg Arg
                                                          -15

CTM MCG AWT GTG GTT GCA TTC GTG GCT CCC GGT GAA TCT CAG CAA GAG    165
Leu Xaa Xaa Val Val Ala Phe Val Ala Pro Gly Glu Ser Gln Gln Glu
```

```
            -10                      -5                   1

GAA CCA CCA ACT GAC AAT CAG GAT ATT GAA CCT GGA CAA GAG AGA GAA     213
Glu Pro Pro Thr Asp Asn Gln Asp Ile Glu Pro Gly Gln Glu Arg Glu
     5                  10                      15

GGA ACA CCT CCG ATC GAA GAA CGT AAA GTA GAA GGT GAT TGC CAG GAA     261
Gly Thr Pro Pro Ile Glu Glu Arg Lys Val Glu Gly Asp Cys Gln Glu
 20                  25                      30                  35

ATG GAT CTG GAA AAG ACT CGG AGT GAG CGT GGA GAT GGC TCT GAT GTA     309
Met Asp Leu Glu Lys Thr Arg Ser Glu Arg Gly Asp Gly Ser Asp Val
                 40                      45                  50

AAA GAG AAG ACT CCA CCT AAT CVT AAG CAT GCT AAG ACT AAA GAA GCA     357
Lys Glu Lys Thr Pro Pro Asn Xaa Lys His Ala Lys Thr Lys Glu Ala
             55                      60                  65

GGA GAT GGG CCA TTG                                                 372
Gly Asp Gly Pro Leu
         70
```

(2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 459 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 299..454
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 278..433
id AA100750
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 160..308
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 133..286
id AA100750
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..159
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100

```
                                   region 1..136
                                   id AA100750
                                   est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 9..355
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 1..347
                             id N68686
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 355..402
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 348..395
                             id N68686
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 400..429
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                             region 394..423
                             id N68686
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 1..241
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 90
                             region 5..245
                             id H24263
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 239..337
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 244..342
                             id H24263
                             est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 13..123
     (C) IDENTIFICATION METHOD: Von Heijne matrix
     (D) OTHER INFORMATION:  score 4.8
                             seq PIVRLLSCPGTVA/KD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

TKTTTTTTAG CA ATG GCG GTT CCC GGC GTG GGG CTC TTG ACC CGT TTG AAC    51
              Met Ala Val Pro Gly Val Gly Leu Leu Thr Arg Leu Asn
```

```
                -35                 -30                 -25

CTG TGT GCC CGG AGA AGA ACT CGA GTC CAG CGG CCT ATC GTC AGG CTT      99
Leu Cys Ala Arg Arg Arg Thr Arg Val Gln Arg Pro Ile Val Arg Leu
                -20                 -15                 -10

TTG AGT TGC CCA GGA ACT GTG GCC AAA GAC CTT AGG AGA GAC GAG CAG     147
Leu Ser Cys Pro Gly Thr Val Ala Lys Asp Leu Arg Arg Asp Glu Gln
             -5                   1                   5

CCT TCA GGG AGC GTG GAG ACA GGC TTT GAA GAC AAG ATT CCC AAA AGG     195
Pro Ser Gly Ser Val Glu Thr Gly Phe Glu Asp Lys Ile Pro Lys Arg
     10                  15                  20

AGA TTC TCG GAG ATG CAA AAT GAA AGA CGA GAA CAG GCA CAG CGG ACT     243
Arg Phe Ser Glu Met Gln Asn Glu Arg Arg Glu Gln Ala Gln Arg Thr
 25                  30                  35                  40

GTT TTA ATA CAT TGC CCA GAG AAA ATC AGT GAA AAC AAG TTT CKK AAA     291
Val Leu Ile His Cys Pro Glu Lys Ile Ser Glu Asn Lys Phe Xaa Lys
                 45                  50                  55

TAT TTA TCC CAA TTT GGA CCT ATT AAT AAT CAT TTC TTC TAT GAA AGC     339
Tyr Leu Ser Gln Phe Gly Pro Ile Asn Asn His Phe Phe Tyr Glu Ser
             60                  65                  70

TTT GGT CTC TAT GCT GTC GTA GAA TTT TGC CAA AAG GAA AGC ATA GGT     387
Phe Gly Leu Tyr Ala Val Val Glu Phe Cys Gln Lys Glu Ser Ile Gly
         75                  80                  85

TCA CTG CAG AAT GGG ACT CAT ACT CCA AGC ACG GCC ATG GAG ACT GCA     435
Ser Leu Gln Asn Gly Thr His Thr Pro Ser Thr Ala Met Glu Thr Ala
     90                  95                 100

ATT CCA TTC AGA TCA CGT TCT TCA                                     459
Ile Pro Phe Arg Ser Arg Ser Ser
105                 110
```

(2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 329 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 109..319
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 95
  region 43..253
  id AA017309

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..124
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 28..59
id AA017309
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(126..250)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..125
id T52392
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 21..200
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq LVILSLKSQTLDA/ET

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

```
AGTAAGTCCC CCCGCCTCGC ATG ATG GCT GCG GTG CCG CCG GGC CTG GAG CCG      53
                      Met Met Ala Ala Val Pro Pro Gly Leu Glu Pro
                      -60                 -55                 -50

TGG AAC CGT GTG AGA ATC CCT AAG GCG GGG AAC CGC AGC GCA GTG ACA       101
Trp Asn Arg Val Arg Ile Pro Lys Ala Gly Asn Arg Ser Ala Val Thr
            -45                 -40                 -35

GTG CAG AAC CCC GGC GCG GCC CTT GAC CTT TGC ATT GCA GCT GTA ATT       149
Val Gln Asn Pro Gly Ala Ala Leu Asp Leu Cys Ile Ala Ala Val Ile
        -30                 -25                 -20

AAA GAA TGC CAT CTC GTC ATA CTG TCG CTG AAG AGC CAA ACC TTA GAT       197
Lys Glu Cys His Leu Val Ile Leu Ser Leu Lys Ser Gln Thr Leu Asp
    -15                 -10                 -5

GCA GAA ACA GAT GTG TTA TGT GCA GTC CTT TAC AGC AAT CAC AAC AGA       245
Ala Glu Thr Asp Val Leu Cys Ala Val Leu Tyr Ser Asn His Asn Arg
    1               5                  10                  15

ATG GGC CGC CAC AAA CCC CAT TTG GCC CTC AAA CAG GTT GAG CAA TGT       293
Met Gly Arg His Lys Pro His Leu Ala Leu Lys Gln Val Glu Gln Cys
                20                  25                  30

TTA AAG CGT TTG ARA AAC ATG AAT TTG GAG GGC GGG                       329
Leu Lys Arg Leu Xaa Asn Met Asn Leu Glu Gly Gly
            35                  40
```

(2) INFORMATION FOR SEQ ID NO: 227:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 385 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 39..385
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..347
id AA023764
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 146..385
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 145..384
id C03036
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 11..80
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 2..71
id C03036
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 39..231
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..193
id R08519
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 232..302
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 193..263
id R08519
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 11..109

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq SLVHLLCQNQVLG/NP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

```
AAGTGGCAAG ATG GCG TCC CTG GAT CGG GTG AAG GTA CTG GTG TTG GGA       49
           Met Ala Ser Leu Asp Arg Val Lys Val Leu Val Leu Gly
                       -30                 -25

GAC TCA GGT GTT GGG AAA TCT TCG TTA GTC CAT CTC CTA TGC CAA AAT      97
Asp Ser Gly Val Gly Lys Ser Ser Leu Val His Leu Leu Cys Gln Asn
-20                 -15                 -10                 -5

CAA GTG CTG GGA AAT CCA TCA TGG ACT GTG GGC TGC TCA GTG GAT GTC     145
Gln Val Leu Gly Asn Pro Ser Trp Thr Val Gly Cys Ser Val Asp Val
                 1               5                  10

AGA GTK CAT GAT TAC AAA GAA GGA ACC CCA GAA GAG AAG ACC TAC TAC     193
Arg Val His Asp Tyr Lys Glu Gly Thr Pro Glu Glu Lys Thr Tyr Tyr
        15                  20                  25

ATA GAA TTA TGG GAT GTT GGA GGC TCT GTG GGC AGT GCC AGC AGC GTG     241
Ile Glu Leu Trp Asp Val Gly Gly Ser Val Gly Ser Ala Ser Ser Val
    30                  35                  40

AAA AGC ACA AGA GCA GTA TTC TAC AAC TCC GTA AAT GGT ATT ATW NYC     289
Lys Ser Thr Arg Ala Val Phe Tyr Asn Ser Val Asn Gly Ile Ile Xaa
45                  50                  55                  60

GTA CAC GAC TTA ACV SAT GGG AAG TCC TCC CAA AAM TTG CGN CGT TGG     337
Val His Asp Leu Thr Xaa Gly Lys Ser Ser Gln Xaa Leu Arg Arg Trp
                65                  70                  75

TCA TTG GAA GCT CTC AAC AGG GAT TTG GTG CCA ACT GGA GTC TTG GTG     385
Ser Leu Glu Ala Leu Asn Arg Asp Leu Val Pro Thr Gly Val Leu Val
            80                  85                  90
```

(2) INFORMATION FOR SEQ ID NO: 228:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 274 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 30..237
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 12..219

id R19497
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 236..270
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 219..253
id R19497
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..238
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..185
id H75597
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 236..270
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 184..218
id H75597
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 60..238
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..179
id H93398
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 236..270
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 178..212
id H93398
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 98..270
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..173
id HUM030E11B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 1..127
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 98
region 118..244
id AA280273
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 50..142
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq WAFSCGTWLPSRA/EW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

```
GCGTCCGCGC CATCAGGCCC GAGATAGCGG CGAGGTCCGC TTTCAGTGT ATG GTT TTC   58
                                                      Met Val Phe
                                                          -30

CCT GCC AAA CGG TTC TGC TTG GTG CCA TCC ATG GAG GGC GTG CGC TGG   106
Pro Ala Lys Arg Phe Cys Leu Val Pro Ser Met Glu Gly Val Arg Trp
            -25                 -20                 -15

GCC TTT TCC TGC GGC ACT TGG CTG CCG AGC CGA GCC GAA TGG CTG CTK   154
Ala Phe Ser Cys Gly Thr Trp Leu Pro Ser Arg Ala Glu Trp Leu Leu
        -10                 -5                  1

RCA GTG CGA TCG ATT CAG CCC GAG GAG AAG GAG CGC ATT GGC CAG TTC   202
Xaa Val Arg Ser Ile Gln Pro Glu Glu Lys Glu Arg Ile Gly Gln Phe
  5                 10                  15                  20

GTC TTT GCC CGG GAC GCT AAG GCA GCC ATG GCT GGT CGT CTG ATG ATA   250
Val Phe Ala Arg Asp Ala Lys Ala Ala Met Ala Gly Arg Leu Met Ile
                25                  30                  35

AGG AAA TTA GTT GCA GAG AAT CGA                                   274
Arg Lys Leu Val Ala Glu Asn Arg
            40
```

(2) INFORMATION FOR SEQ ID NO: 229:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 212 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..208
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 105..223

id HSC13B041
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..99
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 18..115
id HSC13B041
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..208
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 71..189
id T08849
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 19..99
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..81
id T08849
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 19..101
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..83
id H88132
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 90..158
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 71..139
id H88132
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 158..208
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 140..190
id H88132
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 111..208
(C) IDENTIFICATION METHOD: blastn

```
          (D) OTHER INFORMATION:  identity 100
                                  region 92..189
                                  id T33149
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 19..110
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 97
                                  region 1..92
                                  id T33149
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 18..99
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 98
                                  region 1..82
                                  id AA121114
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 158..196
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 100
                                  region 141..179
                                  id AA121114
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 12..89
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4.7
                                  seq LIMQLGSVLLTRC/PF

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:


ACTTTCCCAA G ATG GCG TCG AAG ATA GGT TCG AGA CGG TGG ATG TTG CAG       50
             Met Ala Ser Lys Ile Gly Ser Arg Arg Trp Met Leu Gln
                 -25                 -20                 -15

CTG ATC ATG CAG TTG GGT TCG GTG CTG CTC ACA CGC TGC CCC TTT TGG        98
Leu Ile Met Gln Leu Gly Ser Val Leu Leu Thr Arg Cys Pro Phe Trp
            -10                 -5                  1

GGC TGC TTC AGC CAG CTC ATG CTG TAC GCT GAG AGG GCT GAG GCA CGC       146
Gly Cys Phe Ser Gln Leu Met Leu Tyr Ala Glu Arg Ala Glu Ala Arg
    5                   10                  15

CGG AAG CCC GAC ATC CCA GTG CCT TAC CTG TAT TTC GAC ATG GGG GCA       194
Arg Lys Pro Asp Ile Pro Val Pro Tyr Leu Tyr Phe Asp Met Gly Ala
20                  25                  30                  35

GCC GTG CTG TGC GCG CGG                                               212
Ala Val Leu Cys Ala Arg
                40
```

(2) INFORMATION FOR SEQ ID NO: 230:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 301 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 40..293
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 19..272
id W52056
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 128..220
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LAVDSWWLDPGHA/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230:

```
AAGAACTGCG TCTCGCGACC CAGGCGCGGG TTCCCGGAGG ACAGCCAACA AGCGATGCTG     60

CCGCCGCCGT TTCCTGATTG GTTGTGGGTG GCTACCTCTT CGTTCTGATT GGCCGCTAGT    120

GAGCAAG ATG CTG AGC AAG GGT CTG AAG CGG AAA CGG GAG GAG GAG GAG      169
        Met Leu Ser Lys Gly Leu Lys Arg Lys Arg Glu Glu Glu Glu
            -30                 -25                 -20

GAG AAG GAA CCT CTG GCA GTC GAC TCC TGG TGG CTA GAT CCT GGC CAC      217
Glu Lys Glu Pro Leu Ala Val Asp Ser Trp Trp Leu Asp Pro Gly His
        -15                 -10                 -5

GCA GCG GTG GCA CAG GCA CCC CCG GCC GTG GCC TCT AGC TCC CTC TTT      265
Ala Ala Val Ala Gln Ala Pro Pro Ala Val Ala Ser Ser Ser Leu Phe
    1               5                   10                  15

GAC CTC TCA GTG CTC AAG CTC CAC CAC AGC CGC GGG                      301
Asp Leu Ser Val Leu Lys Leu His His Ser Arg Gly
                20                  25
```

(2) INFORMATION FOR SEQ ID NO: 231:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 380 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..282
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 88..277
id W02951
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 40..93
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 36..89
id W02951
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 347..381
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 345..379
id W02951
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 7..41
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 2..36
id W02951
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 316..347
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 313..344
id W02951
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 283..316
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 94
region 279..312
id W02951
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..305
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 81..293
id N40687
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 12..93
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..82
id N40687
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 305..381
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 292..368
id N40687
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..305
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 80..292
id N44828
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 305..381
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 291..367
id N44828
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 40..93
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 28..81
id N44828
est

(ix) FEATURE:
(A) NAME/KEY: other

(B) LOCATION: 93..381
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 79..367
id R91018
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..93
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..80
id R91018
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 93..305
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 80..292
id W19557
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 13..93
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..81
id W19557
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 305..380
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 291..366
id W19557
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 282..329
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq SLAAALTLHGHWG/LG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231:

AAGGAACGAG ATGGCGGTTC TCTGGAGGCT GAGTGCCGTT TGCGGTGCCC TAGGAGGCCG 60

AGCTCTGTTG CTTCGAACTC CAGTGGTCAG AMCCTGCTCA TATCTCAGCA TTTCTTCAGG 120

ACCGACCTAT CCCAGAATGG TGTGGAGTGC AGCACATACA CTTGTCACCG AGCCACCATT 180

CTGGCTCCAA GGCTGCATCT CTCCACTGGA CTAGCGAGAG GGTTGTCAGT GTTTTGCTCC 240

```
TGGGTCTGCT TCCGGCTGCT TATTTGAATC CTTGCTCTGC G ATG GAC TAT TCC CTG  296
                                            Met Asp Tyr Ser Leu
                                                -15

GCT GCA GCC CTC ACT CTT CAT GGT CAC TGG GGC CTT GGA CAA GTT GTT    344
Ala Ala Ala Leu Thr Leu His Gly His Trp Gly Leu Gly Gln Val Val
    -10                 -5                  1                   5

ACT GAC TAT GTT CAT GGG GAT GCC TTG CAG AAA GCT                    380
Thr Asp Tyr Val His Gly Asp Ala Leu Gln Lys Ala
                10                  15
```

(2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 444 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 138..348
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 128..338
id HUM080D04B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 10..143
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..134
id HUM080D04B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 348..408
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 339..399
id HUM080D04B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 407..445
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94

```
                              region 397..435
                              id HUM080D04B
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 138..274
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:   identity 96
                              region 127..263
                              id H29248
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 11..143
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:   identity 94
                              region 1..133
                              id H29248
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 273..348
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:   identity 98
                              region 263..338
                              id H29248
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 348..387
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:   identity 97
                              region 339..378
                              id H29248
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 382..411
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:   identity 93
                              region 374..403
                              id H29248
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 138..348
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:   identity 97
                              region 128..338
                              id HUM179H02B
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 10..143
```

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..134
id HUM179H02B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 348..397
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 339..388
id HUM179H02B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 407..437
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 396..426
id HUM179H02B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 138..299
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 135..296
id H73551
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..143
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..141
id H73551
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 292..348
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 290..346
id H73551
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 402..441
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 407..446
id H73551
est

(ix) FEATURE:

```
          (A) NAME/KEY: other
          (B) LOCATION: 138..326
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 96
                                  region 94..282
                                  id W68502
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 44..143
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 94
                                  region 1..100
                                  id W68502
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 348..408
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 98
                                  region 306..366
                                  id W68502
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 181..396
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4.7
                                  seq LSLXASYIFGISG/FE

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

AGTTTTCAGG ARATTTGGAA GCTGCCGCAG TAGTTGGAGT CTAAGGACTC GTGACAATCT    60

TCGGGTGCCC TTCGAGAGAA AAGGGGAGGA TGCCACTGGA GTCATCCTCT TCAATGCCAC   120

TATCCTTCCC ATCTBYBYTD RCCCTCRGTA CCACACAATA CTAACCCTTC CCCTNCTCTG   180

ATG TCT TAC ATC ACC TCC CAG GAG ATG AAG TGT ATT CTT CAC TGG TTT      228
Met Ser Tyr Ile Thr Ser Gln Glu Met Lys Cys Ile Leu His Trp Phe
        -70                 -65                 -60

GCC AAT TGG TCA GGT CCC CAG CGT GAA CGT TTC CTA GAG GAC CTG GTA      276
Ala Asn Trp Ser Gly Pro Gln Arg Glu Arg Phe Leu Glu Asp Leu Val
    -55                 -50                 -45

GCT AAG GCA GTG CCA GAA AAA TTA CAA CCA HTG CTG GAT AGT CTG GAG      324
Ala Lys Ala Val Pro Glu Lys Leu Gln Pro Xaa Leu Asp Ser Leu Glu
-40                 -35                 -30                 -25

CAG CTT AGT GTG TCT GGG GCA GAC GAC CAC CTT CTA TCT TTG WGT GCC      372
Gln Leu Ser Val Ser Gly Ala Asp Asp His Leu Leu Ser Leu Xaa Ala
                -20                 -15                 -10

AGC TAC ATC TTT GGG ATC AGT GGT TTC GAG GCT GGG GCT GAG CAG GAG      420
Ser Tyr Ile Phe Gly Ile Ser Gly Phe Glu Ala Gly Ala Glu Gln Glu
            -5                   1               5
```

```
CGC AAT GAA TTT GTC AGA CAG TCG                                   444
Arg Asn Glu Phe Val Arg Gln Ser
     10                  15
```

(2) INFORMATION FOR SEQ ID NO: 233:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 433 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Brain

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 46..406
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 99
  region 6..366
  id W31798
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 55..406
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 99
  region 2..353
  id AA056667
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 68..406
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 99
  region 4..342
  id AA131958
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 35..368
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 99
  region 1..334
  id H10262
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 77..406
- (C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 99
region 1..330
id W95790
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 200..427
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LIVYLWVVSFIAS/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

```
AAGACGAGGT CATGAATCAT GTGACGGTGG CTTGAGGAGG AACCTGTCTT TAAAGCTGTC    60

CCTGAAGTGA CAGCGGAGAG AACCAGGCAG CCCAGAAACC CCAGGCGTGG AGATTGATCC   120

TGCGAGAGAA GGGGGTTCAT CATGGCGGAT GACCTAAAGC GATTCTTGTA TAAAAAGTTA   180

CCAAGTGTTG AAGGGCTCC ATG CCA TTG TTG TGT CAG ATA GAG ATG GAG TAC    232
                    Met Pro Leu Leu Cys Gln Ile Glu Met Glu Tyr
                        -75                 -70

CTG TTA TTA AAG TGG CAA ATG ACA ATG CTC CAG AGC ATG CTT TGC GAC    280
Leu Leu Leu Lys Trp Gln Met Thr Met Leu Gln Ser Met Leu Cys Asp
-65                 -60                 -55                 -50

CTG GTT TCT TAT CCA CTT TTG CCC TTG CAA CAG ACC AAG GAA GCA AAC    328
Leu Val Ser Tyr Pro Leu Leu Pro Leu Gln Gln Thr Lys Glu Ala Asn
                -45                 -40                 -35

TTG GAC TTT CCA AAA ATA AAA GTA TCA TCT GTT ACT ATA ACA CCT ACC    376
Leu Asp Phe Pro Lys Ile Lys Val Ser Ser Val Thr Ile Thr Pro Thr
            -30                 -25                 -20

AGG TGG TTC MAT TTA ATC GTT TAC CTT TGG GTG GTG AGT TTC ATA GCC    424
Arg Trp Phe Xaa Leu Ile Val Tyr Leu Trp Val Val Ser Phe Ile Ala
        -15                 -10                  -5

AGC AGC AGT                                                        433
Ser Ser Ser
     1
```

(2) INFORMATION FOR SEQ ID NO: 234:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 245 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 18..158
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 92
                             region 39..179
                             id C15963
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 139..239
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 161..261
                             id C15963
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 17..219
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                             region 22..224
                             id W07092
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(2..239)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 140..377
                             id W72958
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 2..239
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 18..255
                             id W24219
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 2..239
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                             region 16..253
                             id AA040714
                             est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 45..110
     (C) IDENTIFICATION METHOD: Von Heijne matrix
     (D) OTHER INFORMATION:  score 4.7
                             seq SVMGVCLLIPGLA/TA
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:

```
AAAGGACCCA GAAGTAGGGT TTTGGCCTAG GTAACGGGGC AGAG ATG TGG TTC GAG      56
                                                 Met Trp Phe Glu
                                                         -20

ATT CTC CCC GGA CTC TCC GTC ATG GGC GTG TGC TTG TTG ATT CCA GGA      104
Ile Leu Pro Gly Leu Ser Val Met Gly Val Cys Leu Leu Ile Pro Gly
            -15                 -10                 -5

CTG GCT ACT GCG TAC ATC CAC ARG TTC ACT AAC CGG GGC AAG GAA AAA      152
Leu Ala Thr Ala Tyr Ile His Xaa Phe Thr Asn Arg Gly Lys Glu Lys
        1               5                   10

AGG GTT GCT CAT TTT GGG TAT CAC TGG AGT CTG ATG GAA AGA GAT AGG      200
Arg Val Ala His Phe Gly Tyr His Trp Ser Leu Met Glu Arg Asp Arg
 15                 20                  25                  30

CGC ATC TCT GGA GTT GAT CGT TAC TAT GTG TCA AAG GGT CCA GGG          245
Arg Ile Ser Gly Val Asp Arg Tyr Tyr Val Ser Lys Gly Pro Gly
                35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO: 235:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 354 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 204..351
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 162..309
id AA017973
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 204..351
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 181..328
id AA021972
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 204..351
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 93
region 181..328
id AA013987
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 204..351
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 168..315
id AA014054
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 204..351
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 184..331
id W80073
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 205..342
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq LLVSLVLRXPAKS/TR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

```
AGTTTAGCGA CCGGACCCGA AACGGGGAAG TTGTCTTGTG TGGAGAGGTT AGTAAAGCAG    60

CGCGCGCGTC ACCAGAGTCG TTTCTCTTCG GAGTCTTAGG TGATCGAGGG TGTGCCCAGG   120

GGGCGGACTT GTTTGCGCCT CCCGTTCCCT CCCAATTTCC AAACGTGTCA CCCCGGCGCC   180

GACGGCCCTG TGCAGGGGAA GCAG ATG GAG TTC AAG CTG GAG GCT CAT CGC      231
                           Met Glu Phe Lys Leu Glu Ala His Arg
                                -45                 -40

ATC GTC AGC ATC TCT CTG GGC AAG ATC TAC AAC TCG CGG GTC CAG CGC      279
Ile Val Ser Ile Ser Leu Gly Lys Ile Tyr Asn Ser Arg Val Gln Arg
        -35                 -30                 -25

GGC GGC ATC AAG CTG CAT AAG AAC CTC CTG GTC TCG CTG GTG CTG CGC      327
Gly Gly Ile Lys Leu His Lys Asn Leu Leu Val Ser Leu Val Leu Arg
    -20                 -15                 -10

ASG CCC GCC AAG TCT ACC CGA GCG GGG                                  354
Xaa Pro Ala Lys Ser Thr Arg Ala Gly
 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 236:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 420 base pairs

(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 37..215
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..179
id AA146876
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 214..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 179..333
id AA146876
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 370..399
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 334..363
id AA146876
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 49..319
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 15..285
id AA044109
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 371..414
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 338..381
id AA044109
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 339..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 307..336

id AA044109
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 52..362
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 27..337
id H21138
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 372..407
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 347..382
id H21138
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 52..254
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 14..216
id AA150025
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 307..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 271..332
id AA150025
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 253..315
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 216..278
id AA150025
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 370..414
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 332..376
id AA150025
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 59..368
(C) IDENTIFICATION METHOD: blastn

```
        (D) OTHER INFORMATION:  identity 98
                                region 1..310
                                id N28828
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 370..414
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 311..355
                                id N28828
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 94..384
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.6
                                seq IASGLGLXLDCWT/SS

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 236:

AATCTAGCCC CGCCCCAGGC GAGGGCGCCG CACCCACACC GCGCTGCGCA GTTTTGTTCT    60

GCTCCAGCTG TTCGAAGGTG ATCCAGACGC AAG ATG GCT GTC CTC TCT AAG GAA   114
                                     Met Ala Val Leu Ser Lys Glu
                                             -95

TAT GGT TTT GTG CTT CTA ACT GGT GCT GCC AGC TTT ATA ATG GTG GCC   162
Tyr Gly Phe Val Leu Leu Thr Gly Ala Ala Ser Phe Ile Met Val Ala
-90                 -85                 -80                 -75

CAC CTA GCC ATC AAT GTT TCC AAG GCC CGC AAG AAG TAC AAA GTG GAG   210
His Leu Ala Ile Asn Val Ser Lys Ala Arg Lys Lys Tyr Lys Val Glu
                -70                 -65                 -60

TAT CCT ATC ATG TAC AGC ACG GAC CCT GAA AAT GGG CAC ATC TTC AAC   258
Tyr Pro Ile Met Tyr Ser Thr Asp Pro Glu Asn Gly His Ile Phe Asn
            -55                 -50                 -45

TGC ATT CAG CGA GCC CAC CAG AAC ACG TTG GAA GTG TAT CCT CSC TTC   306
Cys Ile Gln Arg Ala His Gln Asn Thr Leu Glu Val Tyr Pro Xaa Phe
        -40                 -35                 -30

TTA TTT TTT CTA GCT GTT GGA GGT GTT TAC CAC CCG CGT ATA GCT TCT   354
Leu Phe Phe Leu Ala Val Gly Gly Val Tyr His Pro Arg Ile Ala Ser
    -25                 -20                 -15

GGC CTG GGC TTG DCN CTG GAT TGT TGG ACG AGT TCT TTA TGC TTA TGG   402
Gly Leu Gly Leu Xaa Leu Asp Cys Trp Thr Ser Ser Leu Cys Leu Trp
-10                 -5                   1               5

CTA TTA CAC GGG CCG GGG                                           420
Leu Leu His Gly Pro Gly
            10

(2) INFORMATION FOR SEQ ID NO: 237:
```

```
(i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 406 base pairs
      (B) TYPE: NUCLEIC ACID
      (C) STRANDEDNESS: DOUBLE
      (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
      (A) ORGANISM: Homo Sapiens
      (F) TISSUE TYPE: Brain

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 28..227
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 99
                              region 1..200
                              id AA074804
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 265..310
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 238..283
                              id AA074804
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 227..263
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 201..237
                              id AA074804
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 352..385
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 94
                              region 328..361
                              id AA074804
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(259..408)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 99
                              region 123..272
                              id N93600
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(85..207)
```

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 325..447
id N93600
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(202..408)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 117..323
id AA074748
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(116..153)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 375..412
id AA074748
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(167..202)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 324..359
id AA074748
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(258..408)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 123..273
id N93603
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(208..251)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 280..323
id N93603
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(163..202)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 329..368
id N93603
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: complement(90..125)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 411..446
id N93603
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 272..397
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq RIPSLPGSPVCWA/WP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

```
AAAAGGAAAG AGGTYSGGAG CGCTCGCGAG ATCTCGGACC ACCCAACCTG AAAGGTGCTT   60

AGGAAGTTGA AAGGCCCAGA GGAGGCCTCC GGGCAAATGG CCGGAGCTGG ACCGACCATG  120

CTGCTACGAG AAGAGAATGG CTGTTGCAGT CGGCGTCAGA GCAGCTCCAG TGCCGGGGAT  180

TCGGACGGAG AGCGCGAGGA CTCGGCGGCT GAGCGCGCCC GACAGCAGCT AGAGGCGCTG  240

CTCAACAAGA CTATGCGCAT TCGCATGACA G ATG GAC GGA CAC TGG TCG GCT     292
                                   Met Asp Gly His Trp Ser Ala
                                           -40

GCT TTC TCT GCA CTG ACC GTG ACT GCA ATG TCA TCC TGG GCT CGG CGC    340
Ala Phe Ser Ala Leu Thr Val Thr Ala Met Ser Ser Trp Ala Arg Arg
-35                 -30                 -25                 -20

AGG AGT TCC TCA AGC CGT CGG ATT CCT TCT CTG CCG GGG AGC CCC GTG    388
Arg Ser Ser Ser Ser Arg Arg Ile Pro Ser Leu Pro Gly Ser Pro Val
                -15                 -10                 -5

TGC TGG GCC TGG CCA TGG                                            406
Cys Trp Ala Trp Pro Trp
            1
```

(2) INFORMATION FOR SEQ ID NO: 238:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Liver

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 56..207

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 20..171
id N41898
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 69..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 38..176
id H69272
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 56..103
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq RLLLRRFLASVIS/RK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:

```
ACTTGACAGG CAGGGAGGGC TAGGCTGTGC ATCCCTCCGC TCGCATTGCA GGGAG ATG      58
                                                            Met

GCT CAG CGA CTT CTT CTG AGG AGG TTC CTG GCC TCT GTC ATC TCC AGG      106
Ala Gln Arg Leu Leu Leu Arg Arg Phe Leu Ala Ser Val Ile Ser Arg
-15                 -10                 -5                   1

AAG CCC TCT CAG GGT CAG TGG CCA CCC CTC ACT TCC AGA GCC CTG CAG      154
Lys Pro Ser Gln Gly Gln Trp Pro Pro Leu Thr Ser Arg Ala Leu Gln
                5                  10                  15

ACC CCA CAA TGC AGT CCT GGT GGC CTG ACT GTA ACA CCC AAC CCA GCC      202
Thr Pro Gln Cys Ser Pro Gly Gly Leu Thr Val Thr Pro Asn Pro Ala
        20                  25                  30

CGG ACG                                                              208
Arg Thr
    35
```

(2) INFORMATION FOR SEQ ID NO: 239:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 400 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 124..343
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 103..322
id H72703
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..135
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 4..115
id H72703
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 357..398
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 336..377
id H72703
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 7..343
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..337
id W68324
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 357..391
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 351..385
id W68324
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 7..134
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..128
id AA054941
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 191..283
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 185..277
id AA054941
est

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 124..191
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 92
                              region 117..184
                              id AA054941
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 361..398
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 92
                              region 360..397
                              id AA054941
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 124..343
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 97..316
                              id AA128297
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 27..134
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 1..108
                              id AA128297
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 357..398
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 330..371
                              id AA128297
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(153..300)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 108..255
                              id H72704
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(291..343)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 64..116
```

```
                         id H72704
                         est

 (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(101..151)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 259..309
                              id H72704
                              est

 (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(357..398)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 9..50
                              id H72704
                              est

 (ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 311..385
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 4.5
                              seq FLLLLEVSHLLLI/IN

 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:

AGACGTGTTC TTCCGGTGGC GGASGGCGGA TTAGCCTTCG CGGGGCAAAA TGGAGCTCGA    60

GGCCATGAGC AGATATACCA GCCCAGTGAA CCCAGCTGTC TTCCCCCATC TGACCGTGGT   120

GCTTTTGGCC ATTGGCATGT TCTTCACCGC CTGGTTCTTC GTTTACGAGG TCACCTCTAC   180

CAAGTACACT CGTGATATCT ATAAAGAGCT CCTCATCTCC TTAGTGGCCT CACTCTTCAT   240

GGGCTTTGGA GTCCTCTTCC TGCTGCTCTG GGTTGGCATC TACGTGTGAG CACCCAAGGG   300

TAACAACCAG ATG GCT TCA CTG AAA CCT GCT TTT GTA AAT TAC TTT TTT      349
           Met Ala Ser Leu Lys Pro Ala Phe Val Asn Tyr Phe Phe
           -25                 -20                 -15

TTA CTG TTG CTG GAA GTG TCC CAC CTG CTG CTC ATA ATA AAT GCA GAA     397
Leu Leu Leu Leu Glu Val Ser His Leu Leu Leu Ile Ile Asn Ala Glu
    -10                 -5                  1

GGG                                                                 400
Gly
  5
```

(2) INFORMATION FOR SEQ ID NO: 240:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 395 base pairs
        (B) TYPE: NUCLEIC ACID
        (C) STRANDEDNESS: DOUBLE
```

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 226..396
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 202..372
id N40054
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 27..162
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 2..137
id N40054
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 158..214
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 134..190
id N40054
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 15..146
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 2..133
id W25483
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 226..305
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 213..292
id W25483
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 157..214
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 144..201
id W25483
est

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 34..157
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 99
                              region 1..124
                              id C17967
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 226..324
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 194..292
                              id C17967
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 157..214
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 125..182
                              id C17967
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 326..387
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 293..354
                              id C17967
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 226..396
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 167..337
                              id N27721
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 61..162
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 1..102
                              id N27721
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 158..214
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 99..155
```

```
                              id N27721
                              est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 50..214
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 99
                                  region 1..165
                                  id T47061
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 226..377
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 99
                                  region 177..328
                                  id T47061
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 156..386
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4.5
                                  seq LFWVIVLTSWITI/FQ

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

AAAAACGTCC ATAACTGAAA GTAGCTAAGG CACCCCAGGC GGAGGAAGTG AGCTCTCCTG    60

GGGCGTGGTT GTTCGTGATC CTTGCATCTG TTACTTAGGG TCAAGGCTTG GGTCTTGCCC   120

CGCAGACCCT TGGGACGACC CGGCCCCAGC GCAST ATG AAC CTG GAG CGA GTG      173
                                      Met Asn Leu Glu Arg Val
                                                  -75

TCC AAT GAG GAG AAA TTG AAC CTG TGC CGG AAG TAC TAC CTG GGG GGG     221
Ser Asn Glu Glu Lys Leu Asn Leu Cys Arg Lys Tyr Tyr Leu Gly Gly
    -70                 -65                 -60

TTT GCT TTC CTG CCT TTT CTC TGG TTG GTC AAC ATC TTC TGG TTC TTC     269
Phe Ala Phe Leu Pro Phe Leu Trp Leu Val Asn Ile Phe Trp Phe Phe
-55                 -50                 -45                 -40

CGA GAG GCC TTC CTT GTC CCA GCC TAC ACA GAA CAG AGC CAA ATC AAA     317
Arg Glu Ala Phe Leu Val Pro Ala Tyr Thr Glu Gln Ser Gln Ile Lys
                -35                 -30                 -25

GGC TAT GTC TGG CGC TCA GCT GTG GGC TTC CTC TTC TGG GTG ATA GTG     365
Gly Tyr Val Trp Arg Ser Ala Val Gly Phe Leu Phe Trp Val Ile Val
            -20                 -15                 -10

CTC ACC TCC TGG ATC ACC ATC TTC CAG ATC                             395
Leu Thr Ser Trp Ile Thr Ile Phe Gln Ile
        -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 241:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 189 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 80..115
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 170..205
id AA090974
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 73..135
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq AVASSFFCASLFS/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

```
ATTTTTTTCT TGCTCGTGGG CTCGGACGAG TACGGAGCGC CTGCAGGGAC AGCCTGGATA    60

AAGGCTCACT TG ATG GCT CAG TTG GGA GCA GTT GTG GCT GTG GCT TCC AGT   111
              Met Ala Gln Leu Gly Ala Val Val Ala Val Ala Ser Ser
                  -20                 -15                 -10

TTC TTT TGT GCA TCT CTC TTC TCA GCT GTG CAC AAG ATA GAA GAG GGA     159
Phe Phe Cys Ala Ser Leu Phe Ser Ala Val His Lys Ile Glu Glu Gly
             -5                   1               5

CAT ATT GGG GTA TAT TAC AGA GGC GGT GTG                             189
His Ile Gly Val Tyr Tyr Arg Gly Gly Val
     10                  15
```

(2) INFORMATION FOR SEQ ID NO: 242:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 313 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

```
        (D) DEVELOPMENTAL STAGE: Fetal
        (F) TISSUE TYPE: kidney

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 62..308
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 95
                                region 16..262
                                id AA044042
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 46..78
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 90
                                region 1..33
                                id AA044042
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 75..308
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 6..239
                                id AA127902
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 93..308
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 1..216
                                id AA056679
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(104..308)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 246..450
                                id W93399
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 126..308
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 2..184
                                id H39528
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 122..196
        (C) IDENTIFICATION METHOD: Von Heijne matrix
```

(D) OTHER INFORMATION: score 4.4
seq LVFMVPLVGLIHL/GW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

```
GCGAAGGTTG TCGGGATCCG CGGCAGCAGC GGCTGCTTGA GATCTGTTTC TGGGGCCTCT   60

GGCGGTGGCG GCCTGGGGCG GCGCGACGGC TGGTGCGCAG GTACACTGAT GCTGAAGTAC  120

T ATG AGC CTT CGG AAC TTG TGG AGA GAC TAC AAA GTT TTG GTT TTT ATG  169
  Met Ser Leu Arg Asn Leu Trp Arg Asp Tyr Lys Val Leu Val Phe Met
  -25                 -20                 -15                 -10

GTC CCT TTA GTT GGG CTC ATA CAT TTG GGG TGG TAC AGA ATC AAA AGC    217
Val Pro Leu Val Gly Leu Ile His Leu Gly Trp Tyr Arg Ile Lys Ser
                 -5                   1               5

AGC CCT GTT TTC CAA ATA CCT AAA AAC GAC GAC ATT CCT GAG CAA GAT    265
Ser Pro Val Phe Gln Ile Pro Lys Asn Asp Asp Ile Pro Glu Gln Asp
         10                  15                  20

AGT CTG GGA CTT TCA AAT CTT CAG AAG AGC CAA ATC CAG GGG ATA CTG    313
Ser Leu Gly Leu Ser Asn Leu Gln Lys Ser Gln Ile Gln Gly Ile Leu
     25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 243:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 415 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 57..306
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 33..282
id AA088487
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 341..409
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq VFCLLISIPTPSA/HL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

```
AGTCGTTGCC ATSGATCCTG GGGACGACTG GCTGGTGGAA TCCTTGCGCT TGTAAATCGT    60

ACCAGGATTT CTATGCATTC GACCTGTCAG GAGCCACTCG AGTCCTTGAA TGGATTGATG   120

ACAAAGGAGT CTTTGTTGCT GGCTATGAAA GCCTGAAAAA GAATGAAATT CTTCATCTGA   180

AATTACCTCT CAGACTTTCT GTAAAGGAAA ACAAGGGCTT ATTCCCAGAA AGAGATTTCA   240

AAGTGCGCCA TGGAGGATTT TCAGACAGGT CTATCTTTGA TCTAAAGCAT GTGCCACATA   300

CCAGGTATGG TCAATTTTGT GATCCAGCCA TCCACACAGG ATG GGA TGG GAT GGC     355
                                            Met Gly Trp Asp Gly
                                                        -20

TGC AAA TGC CTG GGG GTA TTC TGC CTC CTC ATC TCC ATT CCC ACC CCC     403
Cys Lys Cys Leu Gly Val Phe Cys Leu Leu Ile Ser Ile Pro Thr Pro
            -15                 -10                 -5

TCA GCA CAC CTG                                                     415
Ser Ala His Leu
         1
```

(2) INFORMATION FOR SEQ ID NO: 244:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 458 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (D) DEVELOPMENTAL STAGE: Fetal
- (F) TISSUE TYPE: kidney

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 156..451
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 99
  region 122..417
  id AA085629
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 44..144
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 94
  region 14..114
  id AA085629
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 156..259
- (C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 99
region 134..237
id AA132309
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..144
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 29..126
id AA132309
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 274..314
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 254..294
id AA132309
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..144
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 15..112
id H35088
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 156..345
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 63..252
id HUML11153
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 12..365
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq ILAHRLGLIPIHA/DP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

```
AGAGATTGAA G ATG GCG GCT TCT CAG GCG GTG GAG GAA ATG CGG ACC GCG      50
             Met Ala Ala Ser Gln Ala Val Glu Glu Met Arg Thr Ala
                     -115                -110

TGG TTC TGG GGG AGT TTG GGG TTC GCA ATG TCC ATA CTA CTG ACT TTC       98
Trp Phe Trp Gly Ser Leu Gly Phe Ala Met Ser Ile Leu Leu Thr Phe
-105                -100                -95                 -90

CCG GTA ACT ATT CCG GTT ATG ATG ATG CCT GGG ACC AGG RMC GGY TTC      146
Pro Val Thr Ile Pro Val Met Met Met Pro Gly Thr Arg Xaa Gly Phe
```

```
                    -85                 -80                 -75

GAA GRA AGA AWT TTC CGT GTG GAT GTA GTA CAC ATG GAT GAA AAC TCA      194
Glu Xaa Arg Xaa Phe Arg Val Asp Val Val His Met Asp Glu Asn Ser
            -70                 -65                 -60

CTG GAG TTT GAC ATG GTG GGA ATT GAC GCA GCC ATT GCC AAT GCT TTT      242
Leu Glu Phe Asp Met Val Gly Ile Asp Ala Ala Ile Ala Asn Ala Phe
        -55                 -50                 -45

CGA CGA ATT CTG CTA GCT GAG GTG CCA ACT ATG GCT GTG GAG AAG GTC      290
Arg Arg Ile Leu Leu Ala Glu Val Pro Thr Met Ala Val Glu Lys Val
    -40                 -35                 -30

CTG GTG TAC AAT AAT ACA TCC ATT GTT CAG GAT GAG ATT CTT GCT CAC      338
Leu Val Tyr Asn Asn Thr Ser Ile Val Gln Asp Glu Ile Leu Ala His
-25                 -20                 -15                 -10

CGT CTG GGG CTC ATT CCC ATT CAT GCT GAT CCC CGT CTT TTT GAG TAT      386
Arg Leu Gly Leu Ile Pro Ile His Ala Asp Pro Arg Leu Phe Glu Tyr
                -5                  1                   5

CGG AAC CAA GGA GAT GAA GAA GGC ACA GAG ATA GAT ACT CTA CAG TTT      434
Arg Asn Gln Gly Asp Glu Glu Gly Thr Glu Ile Asp Thr Leu Gln Phe
        10                  15                  20

CGT CTC CAG GTC AGA TGC ACT CGG                                      458
Arg Leu Gln Val Arg Cys Thr Arg
    25                  30
```

(2) INFORMATION FOR SEQ ID NO: 245:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 383 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Muscle

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 61..188
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 97
  region 45..172
  id AA156837
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 252..334
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 90
  region 234..316

id AA156837
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 189..256
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 172..239
id AA156837
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 16..64
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..49
id AA156837
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 15..220
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..206
id AA196478
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 252..334
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 236..318
id AA196478
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 222..256
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 207..241
id AA196478
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 61..226
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 44..209
id AA181144
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 252..334
(C) IDENTIFICATION METHOD: blastn

```
        (D) OTHER INFORMATION:  identity 90
                                region 235..317
                                id AA181144
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 17..64
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 1..48
                                id AA181144
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 225..256
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 209..240
                                id AA181144
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 185..334
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 94
                                region 169..318
                                id AA228369
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 61..184
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 46..169
                                id AA228369
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 22..64
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 95
                                region 8..50
                                id AA228369
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 15..219
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 1..205
                                id W04828
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
```

```
          (B) LOCATION: 252..334
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 90
                                  region 236..318
                                  id W04828
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 341..380
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 100
                                  region 323..362
                                  id W04828
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 221..256
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 97
                                  region 206..241
                                  id W04828
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 12..242
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4.4
                                  seq FEARIALLPLLQA/ET

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 245:

ATACTGCGAG T ATG GCG GCG TCA AAG GTG AAA CAG GAC ATG CCT CCG CCG      50
             Met Ala Ala Ser Lys Val Lys Gln Asp Met Pro Pro Pro
                     -75                 -70                 -65

GGG GGC TAT GGG CCC ATC GAC TAC AAA CGG AAC TTR CCG CGT CGA GGA       98
Gly Gly Tyr Gly Pro Ile Asp Tyr Lys Arg Asn Leu Pro Arg Arg Gly
                -60                 -55                 -50

CTG TCG GGC TAC AGC ATG CTG GCC ATA GGG ATT GGA ACC CTG ATC TAC      146
Leu Ser Gly Tyr Ser Met Leu Ala Ile Gly Ile Gly Thr Leu Ile Tyr
            -45                 -40                 -35

GGG CAC TGG AGC ATA ATG AAG TGG AAC CGT GAG CGC AGG CGC CTA CAA      194
Gly His Trp Ser Ile Met Lys Trp Asn Arg Glu Arg Arg Arg Leu Gln
        -30                 -25                 -20

ATC GAG GAC TTC GAG GCT CGC ATC GCG CTG TTG CCA CTG TTA CAG GCA      242
Ile Glu Asp Phe Glu Ala Arg Ile Ala Leu Leu Pro Leu Leu Gln Ala
    -15                 -10                  -5

GAA ACC GAC CGG ARG ACC TTG CAG ATG CTT CGG GAG AAC CTG GAG GAG      290
Glu Thr Asp Arg Xaa Thr Leu Gln Met Leu Arg Glu Asn Leu Glu Glu
  1               5                  10                  15

GAG GCC ATC ATC ATG MAG GAC GTS CYC GAC TGG AAS GTG GGG RAA KVV      338
Glu Ala Ile Ile Met Xaa Asp Val Xaa Asp Trp Xaa Val Gly Xaa Xaa
```

20 25 30

GHT GTT CCA CAC AAC CCG CTG GGT GCC CCC CTT GAT CGG GGA GCT 383
Xaa Val Pro His Asn Pro Leu Gly Ala Pro Leu Asp Arg Gly Ala
35 40 45

(2) INFORMATION FOR SEQ ID NO: 246:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 310 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 58..271
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 54..267
id AA027968
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 105..289
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 94..278
id N90497
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 10..108
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..99
id N90497
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 63..307
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 21..265
id HSC0SD021
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 41..299

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..259
id T31694
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 116..274
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 61..219
id R38457
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..107
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..53
id R38457
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 273..307
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 219..253
id R38457
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 164..289
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq VLFFTGWWIIIDA/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 246:

```
AATGCGCGAC TGAGCCGGGT GGATGGTACT GCTGCATCCG GGTGTCTGGA GGCTGTGGCC    60

GTTTTGTTTT CTTGGCTAAA ATCGGGGGAG TGAGGCGGGC CGGCGCGGCG CGACACCGGG   120

CTCCGGAACC ACTGCACGAC GGGGCTGGAC TGACCTGAAA AAA ATG TCT GGA TTT     175
                                                Met Ser Gly Phe
                                                        -40

CTA GAG GGC TTG AGA TGC TCA GAA TGC ATT GAC TGG GGG GAA AAG CGC     223
Leu Glu Gly Leu Arg Cys Ser Glu Cys Ile Asp Trp Gly Glu Lys Arg
        -35                 -30                 -25

AAT ACT ATT GCT TCC ATT GCT GCT GGT GTA CTA TTT TTT ACA GGC TGG     271
Asn Thr Ile Ala Ser Ile Ala Ala Gly Val Leu Phe Phe Thr Gly Trp
        -20                 -15                 -10

TGG ATT ATC ATA GAT GCA GCT GTT ATT TAT CCC ACC CGG                 310
Trp Ile Ile Ile Asp Ala Ala Val Ile Tyr Pro Thr Arg
```

-5 1 5

(2) INFORMATION FOR SEQ ID NO: 247:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 398 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 101..386
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 73..358
id AA133050
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 71..100
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 42..71
id AA133050
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 168..313
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 145..290
id AA159550
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 71..169
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 47..145
id AA159550
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 339..394
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 319..374
id AA159550

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..68
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 10..45
id AA159550
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 225..356
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq LVFLTFLSIPSFV/GL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

```
AAGGTGCTCG TCATGCGCAA TGTGGCGCTG CGGCGGGCGG CAGGGCCTGT GTGTGCTGAG    60

GCGGCTGAGC GGCGGACATG CACACCACAG AGCGTGGCGA TGGAACAGTA ACCGGGCTTG   120

TGAGAGGGCT CTGCAGTATA AACTAGGAGA CAAGATCCAT GGATTCACCG TAAACCAGGT   180

GACATCTGTT CCCGAGCTGT TCCTGACTGC AGTGAAGCTC ACCC ATG ATG ACA CAG    236
                                                 Met Met Thr Gln

GAG CCA GGT ATT TAC ACC TGG CCA GAG AAA ACA CGA ATA ATC TGT TCA     284
Glu Pro Gly Ile Tyr Thr Trp Pro Glu Lys Thr Arg Ile Ile Cys Ser
-40                 -35                 -30                 -25

GCG TGC AGT TCC GTA CCA CTC CCA TGG ACA GTA CTG GTG TTC CTC ACA     332
Ala Cys Ser Ser Val Pro Leu Pro Trp Thr Val Leu Val Phe Leu Thr
                -20                 -15                 -10

TTC TTG AGC ATA CCG TCC TTT GTG GGT CTC AGA AAT ATC CGT GCA GAG     380
Phe Leu Ser Ile Pro Ser Phe Val Gly Leu Arg Asn Ile Arg Ala Glu
             -5                   1                5

ACC TTT CTT CAA AAT GTT                                             398
Thr Phe Leu Gln Asn Val
     10
```

(2) INFORMATION FOR SEQ ID NO: 248:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 458 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(53..194)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 95
                              region 444..585
                              id AA161193
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(227..324)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 94
                              region 311..408
                              id AA161193
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(328..406)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 230..308
                              id AA161193
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(408..446)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 92
                              region 188..226
                              id AA161193
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(328..406)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 75..153
                              id R06283
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(275..324)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 156..205
                              id R06283
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: complement(408..446)
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 92
                              region 33..71
```

id R06283
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 328..384
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 286..342
id AA152388
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 131..183
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 89..141
id AA152388
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 283..324
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 242..283
id AA152388
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 42..85
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 1..44
id AA152388
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 351..406
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 312..367
id AA159107
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 408..445
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 371..408
id AA159107
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 193..225
(C) IDENTIFICATION METHOD: blastn

```
        (D) OTHER INFORMATION:  identity 93
                                region 166..198
                                id AA159107
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(171..324)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 90
                                region 313..466
                                id AA152366
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(328..406)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 232..310
                                id AA152366
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(408..446)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 92
                                region 190..228
                                id AA152366
                                est

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: 39..80
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 4.4
                                seq FLTALLWRGRIPG/RQ

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:


AGCGGAGACG CAGAGTCTTG AGCAGCGCGN CAGGCACC ATG TTC CTG ACT GCG CTC     56
                                          Met Phe Leu Thr Ala Leu
                                                          -10

CTC TGG CGC GGC CGC ATT CCC GGC CGT CAG TGG ATC GGG AAG CAC CGG      104
Leu Trp Arg Gly Arg Ile Pro Gly Arg Gln Trp Ile Gly Lys His Arg
             -5                  1                   5

CGG CCG CGG TTC GTG TCG TTG CGC GCC AAG CAG AAC ATG ATC CGC CGC      152
Arg Pro Arg Phe Val Ser Leu Arg Ala Lys Gln Asn Met Ile Arg Arg
     10                  15                  20

CTG GAG ATC GAG GCG GAG AAC CAT TAC TGG CTG AGC ATG CCC TAC ATG      200
Leu Glu Ile Glu Ala Glu Asn His Tyr Trp Leu Ser Met Pro Tyr Met
 25                  30                  35                  40

ACC CGG GAG CAG GAG CGC GGC CAC GCC SSG TTG CGC AGG AGG GAG GCC      248
Thr Arg Glu Gln Glu Arg Gly His Ala Xaa Leu Arg Arg Arg Glu Ala
                 45                  50                  55
```

```
TTC GAG GCS ATA AAG GCG GCC GCC ACT TCC AAG TTC CCC CCG CAT AGA    296
Phe Glu Ala Ile Lys Ala Ala Ala Thr Ser Lys Phe Pro Pro His Arg
             60                  65                  70

TTC ATT GCG GAC CAG CTC GAC CAT CTC AVK VGT CAC CAA GAA ATG GTC    344
Phe Ile Ala Asp Gln Leu Asp His Leu Xaa Xaa His Gln Glu Met Val
         75                  80                  85

CTA ATC CTG AGT CGT CAC CCT TGG ATT TTA TGG ATC ACG GAG CTG ACC    392
Leu Ile Leu Ser Arg His Pro Trp Ile Leu Trp Ile Thr Glu Leu Thr
     90                  95                 100

ATC TTT ACC TGG TCT GGA CTG AAA AAC TGT AGC TTG TGT GAA AAT GAG    440
Ile Phe Thr Trp Ser Gly Leu Lys Asn Cys Ser Leu Cys Glu Asn Glu
105                 110                 115                 120

CTT TGG ACC AGT CTT TAT                                            458
Leu Trp Thr Ser Leu Tyr
                125
```

(2) INFORMATION FOR SEQ ID NO: 249:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 398 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 20..400
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..381
id W56872
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 27..317
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..291
id W31727
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 22..375
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99

```
                                region 1..354
                                id W16469
                                est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 45..400
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 99
                                  region 1..356
                                  id N31028
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 22..375
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 98
                                  region 1..354
                                  id W16470
                                  est

     (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 120..389
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4.3
                                  seq TCLTACWTALCCC/CL

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

AACTTGCTCT GAGACAGGTG CGGCAAGTCT ACTGCGGGCT GGTCCGGGCT CCTCAGGTTC     60

AGACCCGACC GTTATCCAGT CGGTTCGTGG AGAGGAGAGG TGSACTTTAC AGGTCCCCG     119

ATG AAC CAA GAG AAC CCT CCA CCA TAT CCA GGC CCT GGT CCA ACG GCC      167
Met Asn Gln Glu Asn Pro Pro Pro Tyr Pro Gly Pro Gly Pro Thr Ala
-90                 -85                 -80                 -75

CCA TAC CCA CCT TAT CCA CCA CAA CCA ATG GGT CCA GGA CHT ATG GGG      215
Pro Tyr Pro Pro Tyr Pro Pro Gln Pro Met Gly Pro Gly Xaa Met Gly
                -70                 -65                 -60

GGA CCC TAC CCA CCT CCT CAA GGG TAC CCC TAC CAA GGA TAC CCA CAG      263
Gly Pro Tyr Pro Pro Pro Gln Gly Tyr Pro Tyr Gln Gly Tyr Pro Gln
            -55                 -50                 -45

TAC GGC TGG CAG GGT GGA CCT CAG GAG CCT CCT AAA ACC ACA GTG TAT      311
Tyr Gly Trp Gln Gly Gly Pro Gln Glu Pro Pro Lys Thr Thr Val Tyr
        -40                 -35                 -30

GTG GTA GAA GAC CAA AGA AGA GAT GAG CTA GGA CCA TCC ACC TGC CTC      359
Val Val Glu Asp Gln Arg Arg Asp Glu Leu Gly Pro Ser Thr Cys Leu
    -25                 -20                 -15

ACA GCC TGC TGG ACG GCT CTC TGT TGC TGC TGT CTC TGG                  398
Thr Ala Cys Trp Thr Ala Leu Cys Cys Cys Cys Leu Trp
-10                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 250:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..331
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 56..332
id AA022276
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..57
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 4..59
id AA022276
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 329..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 329..368
id AA022276
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..284
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 44..273
id W87295
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 284..331
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 274..321
id W87295
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 12..57
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 2..47
id W87295
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 329..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 318..357
id W87295
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 68..331
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..264
id W01758
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 329..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 261..300
id W01758
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 60..259
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 50..249
id W57829
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 12..58
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 3..49
id W57829
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 22..235
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..214
id HUM417E03B

est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 11..172
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LIVWLLVKSFSES/GI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

```
AAGTTCCGCC ATG GCC TCC TTG GAA GTC AGT CGT AGT CCT CGC AGG TCT       49
           Met Ala Ser Leu Glu Val Ser Arg Ser Pro Arg Arg Ser
                           -50                 -45

CGG CGG GAG CTG GAA GTG CGC AGT CCA CGA CAG AAC AAA CAT TCG GTG      97
Arg Arg Glu Leu Glu Val Arg Ser Pro Arg Gln Asn Lys His Ser Val
    -40                 -35                 -30

CTT TTA CCT ACC TAC AAC GAG CGC GAR GAA CTG CCG CTC ATC GTG TGG     145
Leu Leu Pro Thr Tyr Asn Glu Arg Glu Glu Leu Pro Leu Ile Val Trp
-25                 -20                 -15                 -10

CTG CTG GTG AAA AGC TTC TCC GAG AGT GGA ATC AAC TAT GAA ATT ATA     193
Leu Leu Val Lys Ser Phe Ser Glu Ser Gly Ile Asn Tyr Glu Ile Ile
                -5                   1               5

ATC ATA GAT GAT GGA AGC CCA GAT GGA ACA AGG GAT GTT GCT GAA CAG     241
Ile Ile Asp Asp Gly Ser Pro Asp Gly Thr Arg Asp Val Ala Glu Gln
        10                  15                  20

TTG GAG AAG ATC TAT GGG TCA GAC AGA ATT CTT CTA AGA CCA CGA GAG     289
Leu Glu Lys Ile Tyr Gly Ser Asp Arg Ile Leu Leu Arg Pro Arg Glu
    25                  30                  35

AAA AAG TTG GGA CTA GGA ACT GCA TAT ATT CAT GGA ATG RAA ACA TGC     337
Lys Lys Leu Gly Leu Gly Thr Ala Tyr Ile His Gly Met Xaa Thr Cys
40                  45                  50                  55

CAC AGG RAA CTA CAT CAT TAT TAT GGA TGC                             367
His Arg Xaa Leu His His Tyr Tyr Gly Cys
                60                  65
```

(2) INFORMATION FOR SEQ ID NO: 251:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 407 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 70..408
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 61..399
id AA114853
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 19..68
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 11..60
id AA114853
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 18..402
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 7..391
id W23545
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 70..409
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 42..381
id AA069652
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 28..68
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..41
id AA069652
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 18..343
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 8..333
id AA084987
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 63..409
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..347
id AA101916
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 303..344
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq CPTCLCAPSXXWG/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

```
ATCCGGTGCA CGCGAGTSTT CTGAAACGTC AGCTGCGCTC CCCTAGGAGT GCTGAGCCCG    60

CGGAACCGCA GCCATGACTG AGGCTGATGT GAATCCAAAG GYCTATCCCC TTGCCGATGC   120

CCACCTCACC AAGAAGCTAC TGGACCTCGT TCAGCAGTCA TGTAACTATA AGCAGCTTCG   180

GAAAGGWGCC AATGAGGCCA CCAAAACCCT CAACAGGGGC ATCTCTGAGT TCATCGTGAT   240

GGCTGCAGAC GCCGAGCCAC TGGAGATCAT TCTGCACCTG CCGCTGCTGT GTGAAGACAA   300

GA ATG TGC CCT ACG TGT TTG TGC GCT CCA AGC AVN SCC TGG GGA GAG      347
   Met Cys Pro Thr Cys Leu Cys Ala Pro Ser Xaa Xaa Trp Gly Glu
                   -10                  -5                   1

CCT GTG GGG TCT CCA GGC CTG TCA TCG CCT GTT CTG TCA CCA TCA AAG     395
Pro Val Gly Ser Pro Gly Leu Ser Ser Pro Val Leu Ser Pro Ser Lys
             5                  10                  15

AAG GCT CGC AGC                                                     407
Lys Ala Arg Ser
        20
```

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 168 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 43..168
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 34..159
id N52621
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 8..38

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..31
id N52621
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 21..168
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 5..152
id AA157163
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 10..66
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq AVAASAASGQAEG/KK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:

```
ACTTCTAAG ATG GCT GCC GCT ACC GGT GCG GTG GCA GCC TCG GCC GCC TCG     51
          Met Ala Ala Ala Thr Gly Ala Val Ala Ala Ser Ala Ala Ser
                          -15                     -10

GGT CAG GCG GAA GGT AAA AAG ATC ACC GAT CTG CGG GTC ATC GAT CTG       99
Gly Gln Ala Glu Gly Lys Lys Ile Thr Asp Leu Arg Val Ile Asp Leu
 -5                  1                   5                  10

AAG TCC GAG CTG AAG CGG CGG AAC TTA GAC ATC ACC GGA GTC AAG ACC      147
Lys Ser Glu Leu Lys Arg Arg Asn Leu Asp Ile Thr Gly Val Lys Thr
            15                  20                  25

GTG CTC ATC TCC CGA CTA AGG                                          168
Val Leu Ile Ser Arg Leu Arg
        30
```

(2) INFORMATION FOR SEQ ID NO: 253:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 433 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 132..343

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 95..306
id AA102280
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 37..139
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..103
id AA102280
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 340..433
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 304..397
id AA102280
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 132..433
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 87..388
id R13711
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 71..139
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 27..95
id R13711
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 132..401
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 87..356
id R61022
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 71..139
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 27..95
id R61022
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 132..389
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 82..339
id N44705
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 50..139
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..90
id N44705
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 387..433
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 338..384
id N44705
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 126..433
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 75..382
id H29689
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 23..73
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq SLLXRVSVTAVAA/LS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

```
ATTCCTCCTG CCCGTAGTAG CC ATG GCG GCC ATG AGT TTG TTG CKG CGG GTT     52
                         Met Ala Ala Met Ser Leu Leu Xaa Arg Val
                                 -15                 -10

TCG GTT ACT GCG GTG GCA GCT CTG TCT GGC CGG CCC CTT GGC ACY NGC     100
Ser Val Thr Ala Val Ala Ala Leu Ser Gly Arg Pro Leu Gly Thr Xaa
        -5                   1               5

CTC GGA TTT GGG GGC TTC CTC ACT CGT GGC TTT CCG AAG GCT GCT GCT     148
Leu Gly Phe Gly Gly Phe Leu Thr Arg Gly Phe Pro Lys Ala Ala Ala
 10                  15                  20                  25

CCT GTT CGA CAC AGT GGA GAC CAT GGG AAA AGA CTA TTT GTC ATC AGA     196
Pro Val Arg His Ser Gly Asp His Gly Lys Arg Leu Phe Val Ile Arg
                 30                  35                  40
```

```
CCT TCT AGA TTC TAT GAC AGG CGT TTT TTG AAG TTA TTG AGA TTC TAC      244
Pro Ser Arg Phe Tyr Asp Arg Arg Phe Leu Lys Leu Leu Arg Phe Tyr
             45                  50                  55

ATT GCA TTG ACT GGG ATT CCA GTA GCA WTT TTC ATA ACT CTG GTG AAT      292
Ile Ala Leu Thr Gly Ile Pro Val Ala Xaa Phe Ile Thr Leu Val Asn
         60                  65                  70

GTA TTC ATT GGT CAA GCT GAA CTA GCA GAA ATT CCA GAA GGC TAT GTC      340
Val Phe Ile Gly Gln Ala Glu Leu Ala Glu Ile Pro Glu Gly Tyr Val
     75                  80                  85

CCA GAA CAC TGG GAA TAT TAT AAG CAT CCC ATA TCA AGA TGG ATT GCC      388
Pro Glu His Trp Glu Tyr Tyr Lys His Pro Ile Ser Arg Trp Ile Ala
 90                  95                 100                 105

CGT AAT TTC TAT GAT AGT CCT GMA AAG ATA TAT GAA AGA ACA ATG          433
Arg Asn Phe Tyr Asp Ser Pro Xaa Lys Ile Tyr Glu Arg Thr Met
                110                 115                 120
```

(2) INFORMATION FOR SEQ ID NO: 254:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 453 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 86..452
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 43..409
id W00599
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..96
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 12..54
id W00599
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 108..405
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 107..404
id AA088577

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..100
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 29..96
id AA088577
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 6..41
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 1..36
id AA088577
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 40..189
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 7..156
id R18030
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 188..311
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 156..279
id R18030
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 100..261
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 46..207
id H85485
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 61..135
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq LDLLRGLPRVSLA/NL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

GAGACCACGT GGCCTCCGAG CAGCTCAGGG CGCCCTTGAA AGTTCTTGGA TCTGCGGGTT 60

ATG GCC GGT CCC TTG CAG GGC GGT GGG GCC CGG GCC CTG GAC CTA CTC 108
Met Ala Gly Pro Leu Gln Gly Gly Gly Ala Arg Ala Leu Asp Leu Leu

-25 -20 -15 -10

CGG GGC CTG CCG CGT GTG AGC CTG GCC AAC TTA AAG CCG AAT CCC GGC 156
Arg Gly Leu Pro Arg Val Ser Leu Ala Asn Leu Lys Pro Asn Pro Gly
-5 1 5

TCC AAG AAA CCG GAG AGA AGA CCA AGA GGT CGG AGA AGA GGT AGA AAA 204
Ser Lys Lys Pro Glu Arg Arg Pro Arg Gly Arg Arg Arg Gly Arg Lys
10 15 20

TGT GGC AGA GGC CAT AAA GGA GAA AGG CAA AGA GGA ACC CGG CCC CGC 252
Cys Gly Arg Gly His Lys Gly Glu Arg Gln Arg Gly Thr Arg Pro Arg
25 30 35

TTG GGC TTT GAG GGA GGC CAG ACT CCA TTT TAC ATC CGA RTC CCA AAA 300
Leu Gly Phe Glu Gly Gly Gln Thr Pro Phe Tyr Ile Arg Xaa Pro Lys
40 45 50 55

TAC GGG TTT AAC GAA GGA CAT AGT TTC AGA CGC CAG TAT AAG CCT TTG 348
Tyr Gly Phe Asn Glu Gly His Ser Phe Arg Arg Gln Tyr Lys Pro Leu
60 65 70

AGT CTC AAT AGA CTG CAG TAT CTT ATT GAT TTG GGT CGT GTT GAT CCT 396
Ser Leu Asn Arg Leu Gln Tyr Leu Ile Asp Leu Gly Arg Val Asp Pro
75 80 85

AGT CAA CCT ATT GAC TTA ACC CAG CTT GTC AAT GGG AGA GGT GTG ACC 444
Ser Gln Pro Ile Asp Leu Thr Gln Leu Val Asn Gly Arg Gly Val Thr
90 95 100

ATC GCG CCG 453
Ile Ala Pro
105

(2) INFORMATION FOR SEQ ID NO: 255:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 425 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..135
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..103
id T11164
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 133..223
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 102..192
id T11164
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 18..140
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq GILILWIIRLLFS/KT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

```
AAAGGAAGCG GCTAACT ATG GCG ACC GCC ACG GAG CAG TGG GTT CTG GTG        50
                   Met Ala Thr Ala Thr Glu Gln Trp Val Leu Val
                       -40                 -35

GAG ATG GTA CAG GCG CTT TAC GAG GCT CCT GCT TAC CAT CTT ATT TTG        98
Glu Met Val Gln Ala Leu Tyr Glu Ala Pro Ala Tyr His Leu Ile Leu
-30                 -25                 -20                 -15

GAA GGG ATT CTG ATC CTC TGG ATA ATC AGA CTT CTT TTC TCT AAG ACT       146
Glu Gly Ile Leu Ile Leu Trp Ile Ile Arg Leu Leu Phe Ser Lys Thr
                -10                 -5                   1

TAC AAA TTA CAA GAA CGA TCT GAT CTT ACA GTC AAG GAA AAA GAA GAA       194
Tyr Lys Leu Gln Glu Arg Ser Asp Leu Thr Val Lys Glu Lys Glu Glu
        5                   10                  15

CTG ATT GAA GAG TGG CAA CCA GAA CCT CTT GTT CCT CCT GTC CCA AAA       242
Leu Ile Glu Glu Trp Gln Pro Glu Pro Leu Val Pro Pro Val Pro Lys
    20                  25                  30

GAC CAT CCT GCT CTC AAC TAC AAC ATC GTT TCA GGC CCT CCA AGC CAC       290
Asp His Pro Ala Leu Asn Tyr Asn Ile Val Ser Gly Pro Pro Ser His
 35                 40                  45                  50

AAA ACT GTG GTG AAT GGA AAA GAA TGT ATA AAC TTC GCC TCA TTT AAT       338
Lys Thr Val Val Asn Gly Lys Glu Cys Ile Asn Phe Ala Ser Phe Asn
                55                  60                  65

TTT CTT GGA TTG TTG GAT AAC CCT AGG GTT AAG GCA GCA GCT TTA GCA       386
Phe Leu Gly Leu Leu Asp Asn Pro Arg Val Lys Ala Ala Ala Leu Ala
            70                  75                  80

TCT CTA AAG AAG TAT GGC GTG GGG ACT TGT GGA CCC TGT                   425
Ser Leu Lys Lys Tyr Gly Val Gly Thr Cys Gly Pro Cys
        85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO: 256:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 407 base pairs
(B) TYPE: NUCLEIC ACID

(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 147..328
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 124..305
id W16517
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..149
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 7..123
id W16517
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 326..385
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 302..361
id W16517
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 15..149
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 10..144
id H23328
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 147..276
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 145..274
id H23328
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 276..309
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 275..308
id H23328

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 147..309
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 146..308
id H06320
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 73..149
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 69..145
id H06320
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 5..40
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..36
id H06320
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 146..182
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 222..258
id T62768
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 162..398
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq QGVLFICFTCARS/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

```
AAAAACTGAG GCCTGGGAGC AGGAACCTGT AGGCAGCGCT TGAGGGTAGC GGGATAGCAG    60

CTGCAACGCG CGTGGGAGGC GGGGGCTCTG GGCGGAACAA AAATCACAGG ATGTCAGAGG   120

ATGTTTCCCG GGAAGAACTG GGATAAAGGG GTCCCAGCAC C ATG GAG GAC CCG AAC   176
                                              Met Glu Asp Pro Asn
                                                              -75

CCT GAA GAG AAC ATG ADG CAG CAG GAT TCA CCC AAG GAG AGA AGT CCC      224
Pro Glu Glu Asn Met Xaa Gln Gln Asp Ser Pro Lys Glu Arg Ser Pro
                -70                 -65                 -60

CAG AGC CCA GGA GGC AAC ATC TGC CAC CTG GGG GCC CCG AAG TGC ACC      272
```

```
Gln Ser Pro Gly Gly Asn Ile Cys His Leu Gly Ala Pro Lys Cys Thr
            -55                 -50                 -45

CGC TGC CTC ATC ACC TTC GCA GAT TCC AAG TTS SAG GAG CGT CAC ATG      320
Arg Cys Leu Ile Thr Phe Ala Asp Ser Lys Xaa Xaa Glu Arg His Met
        -40                 -35                 -30

AAG CGG GAG CAC CCA GCG GAC TTC GTG GCC CAG AAG CTG CAG GGG GTC      368
Lys Arg Glu His Pro Ala Asp Phe Val Ala Gln Lys Leu Gln Gly Val
    -25                 -20                 -15

CTC TTC ATC TGC TTC ACC TGC GCC CGC TCC TTC CCC TCT                  407
Leu Phe Ile Cys Phe Thr Cys Ala Arg Ser Phe Pro Ser
-10                 -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 257:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 490 base pairs
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: DOUBLE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: complement(166..452)
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 98
  region 16..302
  id AA062591
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 401..445
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 100
  region 1..45
  id AA158358
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 444..490
- (C) IDENTIFICATION METHOD: blastn
- (D) OTHER INFORMATION: identity 93
  region 43..89
  id AA158358
  est

(ix) FEATURE:
- (A) NAME/KEY: other
- (B) LOCATION: 401..445

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..45
id AA158431
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 444..490
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 43..89
id AA158431
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 65..160
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq RLLSSLLLTMSNN/NP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

```
AAGGATCCTC TACCGGCTTT TCGAGTCAGT GCTGCCGCCG CTGCCCGCGG CTTTGCAGAG   60

CAGG ATG AAT GTG ATA GAC CAC GTG CGG GAC ATG GCG GCC GCG GGG CTG   109
     Met Asn Val Ile Asp His Val Arg Asp Met Ala Ala Ala Gly Leu
             -30                 -25                 -20

CAC TCC AAC GTG CGG CTC CTC AGC AGC TTG TTA CTT ACA ATG AGT AAT    157
His Ser Asn Val Arg Leu Leu Ser Ser Leu Leu Leu Thr Met Ser Asn
    -15                 -10                  -5

AAC AAC GCT GAG TTA TTC TCC CCA CCT CAG AAG TAC CAG CTT TTG GTG    205
Asn Asn Pro Glu Leu Phe Ser Pro Pro Gln Lys Tyr Gln Leu Leu Val
      1               5                  10                  15

TAT CAT GCA GAT TCT CTC TTT CAT GAT AAG GAA TAT CGG AAT GCT GTG    253
Tyr His Ala Asp Ser Leu Phe His Asp Lys Glu Tyr Arg Asn Ala Val
                 20                  25                  30

AGT AAG TAT ACC ATG GCT TTA CAG CAG AAG AAA GCG CTA AGT AAA ACT    301
Ser Lys Tyr Thr Met Ala Leu Gln Gln Lys Lys Ala Leu Ser Lys Thr
             35                  40                  45

TCA AAA GTG AGA CCT TCA ACT GGA AAT TCT GCA TCT ACT CCA CAA AGT    349
Ser Lys Val Arg Pro Ser Thr Gly Asn Ser Ala Ser Thr Pro Gln Ser
         50                  55                  60

CAG TGT CTT CCA TCT GAA ATT GAA GTG AAA TAC AAA ATG GCT GAA TGT    397
Gln Cys Leu Pro Ser Glu Ile Glu Val Lys Tyr Lys Met Ala Glu Cys
     65                  70                  75

TAT ACA ATG CTA AAA CAA GAT AAA GAT GCC ATT GCT ATA CTT GAT GGG    445
Tyr Thr Met Leu Lys Gln Asp Lys Asp Ala Ile Ala Ile Leu Asp Gly
 80                  85                  90                  95

KST CCC TTC AAG ACA AAG AAC TCC CAR AAT AAA CAT GAT GCT GGC        490
Xaa Pro Phe Lys Thr Lys Asn Ser Gln Asn Lys His Asp Ala Gly
```

100 105 110

(2) INFORMATION FOR SEQ ID NO: 258:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 340 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 26..337
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..312
id HSC26F061
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 97..337
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 20..260
id W30546
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 97..283
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 66..252
id H34739
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 125..298
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq LVHHCPTWQWATG/EE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

AGGGTGCTGC CWKCCGGGTG CTGWTGCGAG TCGGTGGCAG CGAGGACATT TTCTGACTCC 60

CTGGCCCCTG ACACGGCTGC ACTTTCCATC CCGTCGCGGG GCCGGCCGCT ACTCCGGCCC 120

CAGG ATG CAG AAT GTG ATT AAT ACT GTG AAG GGA AAG GCA CTG GAA GTG 169
Met Gln Asn Val Ile Asn Thr Val Lys Gly Lys Ala Leu Glu Val

```
            -55                 -50                 -45

GCT GAG TAC CTG ACC CCG GTC CTC AAG GAA TCA AAG TTT AAG GAA ACA     217
Ala Glu Tyr Leu Thr Pro Val Leu Lys Glu Ser Lys Phe Lys Glu Thr
            -40                 -35                 -30

GGT GTA ATT ACC CCA GAA GAG TTT GTG GCA GCT GGA GAT CAC CTA GTC     265
Gly Val Ile Thr Pro Glu Glu Phe Val Ala Ala Gly Asp His Leu Val
        -25                 -20                 -15

CAC CAC TGT CCA ACA TGG CAA TGG GCT ACA GGG GAA GAA TTG AAA GTG     313
His His Cys Pro Thr Trp Gln Trp Ala Thr Gly Glu Glu Leu Lys Val
    -10                 -5                  1               5

AAG GCA TAC CTA CCA ACA GGC AAA TGG                                 340
Lys Ala Tyr Leu Pro Thr Gly Lys Trp
                10
```

(2) INFORMATION FOR SEQ ID NO: 259:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 481 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Colon

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 116..289
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 89..262
id W68068
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 360..428
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 334..402
id W68068
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 286..347
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 260..321
id W68068
est

```
(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 66..114
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 100
                               region 40..88
                               id W68068
                               est

(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 26..69
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 95
                               region 1..44
                               id W68068
                               est

(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 428..465
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 100
                               region 401..438
                               id W68068
                               est

(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 66..289
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 99
                               region 62..285
                               id AA083574
                               est

(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 3..45
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 100
                               region 1..43
                               id AA083574
                               est

(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 401..444
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 93
                               region 401..444
                               id AA083574
                               est

(ix) FEATURE:
       (A) NAME/KEY: other
       (B) LOCATION: 314..347
       (C) IDENTIFICATION METHOD: blastn
       (D) OTHER INFORMATION:  identity 100
                               region 312..345
                               id AA083574
```

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 286..316
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 283..313
id AA083574
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 127..289
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 102..264
id AA001460
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 360..465
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 336..441
id AA001460
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 286..347
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 262..323
id AA001460
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 52..103
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 28..79
id AA001460
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 113..289
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 78..254
id H72445
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 286..347
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100

region 252..313
id H72445
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 66..113
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 32..79
id H72445
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 34..69
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..36
id H72445
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 382..411
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 349..378
id H72445
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 209..472
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq CIQRLPWLLLCRG/IT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

```
AGATCCCGCC TGGGGCCGGC TGAGTGGCAC TTAAGCGGGC CATGCCATGC AACCTTGGGC    60

GCTGCCAACC GTGGGCGAGC TCTGGGTGTG CGGGCGGCCT GGCGCGGCGC TCCGCTGTGT   120

CAGCGTGTTA TGATGCCGTC CCGTACCAAC CTGGCTACTG GAATCCCCAG TAGTAAAGTG   180

AAATATTCAA GGCTCTCCAG CACAGACG ATG GCT ACA TTG ACC TTC AGT TTA      232
                               Met Ala Thr Leu Thr Phe Ser Leu
                                           -85

AGA AAA CCC CTC CAA AGA TCC CTT ATA AGG CCA TCG CAC TTG CCA CTG     280
Arg Lys Pro Leu Gln Arg Ser Leu Ile Arg Pro Ser His Leu Pro Leu
-80                 -75                 -70                 -65

TGC TGT TTT GAT TGG CGC CTT TCT CAT TAT TAT AGG CTC CCT CCT GCT     328
Cys Cys Phe Asp Trp Arg Leu Ser His Tyr Tyr Arg Leu Pro Pro Ala
                -60                 -55                 -50

GTC AGG CTA CAT CAG CAA AGG GGG GGC AGA CCG GGC CGT TCC AGT GCT     376
Val Arg Leu His Gln Gln Arg Gly Gly Arg Pro Gly Arg Ser Ser Ala
```

```
         -45                  -40                  -35

GAT CAT TGG CAT TCT GGT GTT CCT ACC CGG ATT TTA CCA CCT GCG CAT      424
Asp His Trp His Ser Gly Val Pro Thr Arg Ile Leu Pro Pro Ala His
        -30                 -25                  -20

CGC TTA CTA TGC ATC CAA AGG CTA CCG TGG TTA CTC CTA TGC AGG GGG      472
Arg Leu Leu Cys Ile Gln Arg Leu Pro Trp Leu Leu Leu Cys Arg Gly
    -15                 -10                  -5

ATC ACT AGT                                                          481
Ile Thr Ser
  1
```

(2) INFORMATION FOR SEQ ID NO: 260:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 338 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 67..218
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 51..202
id N55991
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 16..74
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..59
id N55991
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 89..231
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..143
id R57473
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 232..339
(C) IDENTIFICATION METHOD: blastn

```
          (D) OTHER INFORMATION:  identity 97
                                  region 143..250
                                  id R57473
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 140..243
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 94
                                  region 195..298
                                  id H79944
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 243..279
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 91
                                  region 299..335
                                  id H79944
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 140..237
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 92
                                  region 97..194
                                  id H70394
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 235..325
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 93
                                  region 193..233
                                  id H70394
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 140..325
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 93
                                  region 80..265
                                  id W31972
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: sig_peptide
          (B) LOCATION: 123..269
          (C) IDENTIFICATION METHOD: Von Heijne matrix
          (D) OTHER INFORMATION:  score 4
                                  seq PSLAAGLLFGSXA/GL

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:

AAATCGCGTT TCCGGAGAGA CCTGGCTGCT GTGTCCCGCG GCTTGCGCTC CGTAGTGGAC     60
```

```
TCCGCGGGCC TTCGGCAGAT GCAGGCCTGG GGTAGTCTCC TTTCTGGACT GAGAAGAGAA  120

GA ATG GAG AAG CCC CTC TTC CCA TTA GTG CCT TTG CAT TGG TTT GGC     167
   Met Glu Lys Pro Leu Phe Pro Leu Val Pro Leu His Trp Phe Gly
                   -45                 -40                 -35

TTT GGC TAC ACA GCA CTG GTT GTT TCT GGT GGG ATC GTT GGC TAT GTA    215
Phe Gly Tyr Thr Ala Leu Val Val Ser Gly Gly Ile Val Gly Tyr Val
                -30                 -25                 -20

AAA ACA GGC AGC GTG CCG TCC CTG GCT GCA GGG CTG CTC TTC GGC AGT    263
Lys Thr Gly Ser Val Pro Ser Leu Ala Ala Gly Leu Leu Phe Gly Ser
            -15                 -10                 -5

VWA GCC GGC CTG GGT GCT TAC CAG CTG TAT CAG GAT CCA AGR AAC GTT    311
Xaa Ala Gly Leu Gly Ala Tyr Gln Leu Tyr Gln Asp Pro Arg Asn Val
        1               5                   10

TGG GGT TTC CTA GCC GCT ACA TCT GTT                                338
Trp Gly Phe Leu Ala Ala Thr Ser Val
15                  20
```

(2) INFORMATION FOR SEQ ID NO: 261:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 302 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 95..241
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 92..238
id R27748
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..90
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..89
id R27748
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 162..298
(C) IDENTIFICATION METHOD: blastn

```
      (D) OTHER INFORMATION:  identity 97
                              region 116..252
                              id T79527
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 2..47
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 11..56
                              id T79527
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 53..90
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 61..98
                              id T79527
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 95..195
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 99
                              region 59..159
                              id R08734
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 194..241
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 93
                              region 159..206
                              id R08734
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 48..90
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 14..56
                              id R08734
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 102..298
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 92
                              region 32..228
                              id H35655
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
```

(B) LOCATION: 102..298
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 108..304
id AA038389
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 108..161
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq VAVGLTIAAAGFA/GR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

```
AGGGGGTTGC GTCGCTCTCT GGTAAAGGCG TGCAGGTGTT GGCCGCGGCC TCTGAGCTGG    60

GATGAGCCGT GCTCCCGGTG GAAGCAAGGG GAGCCCCAGC SGGAGCC ATG GCC AGT     116
                                                    Met Ala Ser

ACA GTG GTA GCA GTT GGA CTG ACC ATT GCT GCT GCA GGA TTT GCA GGC    164
Thr Val Val Ala Val Gly Leu Thr Ile Ala Ala Ala Gly Phe Ala Gly
-15                 -10                 -5                   1

CGT TAC GTT TTG CAA GCC ATG AAG CAT ATG GAG CBT CAA GTA AAA CAA    212
Arg Tyr Val Leu Gln Ala Met Lys His Met Glu Xaa Gln Val Lys Gln
            5                   10                  15

GTT TTT CAA AGC CTA CCA AAA TCT GCC TTC AGT GGT GGC TAT TAT AGA    260
Val Phe Gln Ser Leu Pro Lys Ser Ala Phe Ser Gly Gly Tyr Tyr Arg
        20                  25                  30

GGT BGG TTT GAA CCC ARA ATG RCA AAA CGG GAA GCA GCG GGG            302
Gly Xaa Phe Glu Pro Xaa Met Xaa Lys Arg Glu Ala Ala Gly
    35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO: 262:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 465 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 130..311
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 96..277
id T32007

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..130
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..98
id T32007
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 130..314
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 76..260
id R19207
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 53..130
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..78
id R19207
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 130..314
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 76..260
id R36562
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 53..130
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..78
id R36562
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 130..314
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 76..260
id R59039
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 71..130
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98

region 19..78
id R59039
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 130..314
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 70..254
id T35666
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 59..130
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..72
id T35666
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 136..384
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq AFSFSRLLSQCRP/DC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

```
AAAGTTCTCC TTCCACCTTC CCCCACCCTT CTCTGCCAAC CGCTGTTTCA GCCCCTAGCT    60

GGATTCCAGC CATTGCTGCA GCTGCTCCAC AGGCCTTTTC AGGACCCAAA CAACCGCAGC   120

CGCTGTTCCC CAGGR ATG GTG ATC CGT GTA TAT ATT GCA TCT TCC TCT GGC    171
                 Met Val Ile Arg Val Tyr Ile Ala Ser Ser Ser Gly
                             -80                 -75

TCT ACA GCG ATT AAG AAG AAA CAA CAA GAT GTG CTT GGT TTC CTA GAA     219
Ser Thr Ala Ile Lys Lys Lys Gln Gln Asp Val Leu Gly Phe Leu Glu
    -70                 -65                 -60

GCC AAC AAA ATA GGA TTT GAA GAA AAA GAT ATT GCA GCC AAT GAA GAG     267
Ala Asn Lys Ile Gly Phe Glu Glu Lys Asp Ile Ala Ala Asn Glu Glu
-55                 -50                 -45                 -40

AAT CGG AAG TGG ATG AGA GAA AAT GTA CCT GAA AAT AGT CGA CCA GCG     315
Asn Arg Lys Trp Met Arg Glu Asn Val Pro Glu Asn Ser Arg Pro Ala
                -35                 -30                 -25

GTT CAG GGG CCA CAT GCT TTT CGG TAT AAA GCA TTC TCC TTC TCT AGG     363
Val Gln Gly Pro His Ala Phe Arg Tyr Lys Ala Phe Ser Phe Ser Arg
            -20                 -15                 -10

TTG CTA TCA CAG TGC AGA CCT GAC TGC CTG AAT ATG CTC AGG AGA TTT     411
Leu Leu Ser Gln Cys Arg Pro Asp Cys Leu Asn Met Leu Arg Arg Phe
        -5                   1                   5

AGT CAA TAT TGT CTG TAT TTG GTT ATG GAA AAG GCT CTC CTT TTT TTT     459
```

```
Ser Gln Tyr Cys Leu Tyr Leu Val Met Glu Lys Ala Leu Leu Phe Phe
 10                  15                  20                  25

TTT TTT                                                      465
Phe Phe
```

(2) INFORMATION FOR SEQ ID NO: 263:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 401 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 132..289
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 117..274
id R14800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 15..130
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 1..116
id R14800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 315..368
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 302..355
id R14800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 284..316
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 270..302
id R14800
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 132..330

```
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 113..311
                                id R59757
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 54..130
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 93
                                region 36..112
                                id R59757
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 18..58
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 92
                                region 1..41
                                id R59757
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 132..330
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 94..292
                                id R25047
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 54..130
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 94
                                region 17..93
                                id R25047
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 59..352
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 99
                                region 38..331
                                id R23993
                                est

   (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 163..294
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 92
                                region 163..294
                                id W23811
                                est

   (ix) FEATURE:
```

(A) NAME/KEY: other
(B) LOCATION: 132..194
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 131..193
id W23811
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 305..354
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 308..357
id W23811
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 350..390
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 354..394
id W23811
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 243..368
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq ITSSLFLGRGSVA/SN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

```
AAGAAGCCGG TGGCCGCGCA GGAGGACGGA GCCCTAACCG CAACCCGCGC CGCGCCGCGC    60

CGATTTGATT TGTATCCACT GTCACCAGCA CTGCTCACTT AGGACTTTCT GGATCCAGAC   120

CCAGGCAGCG CACACTGGAC TCTTGAGGAA GAAGGAGACT CTAATTTTGG ATTCCTTGGT   180

GGAGGAAAAT AAAACACTCT GGTCTTGCCG CCAACGATGC AAGTGTGACT GCTGGCGTCT   240

TC ATG AGC TCC AGA GGT CAC AGC ACG CTA CCA AGG ACT CTC ATG GCC      287
   Met Ser Ser Arg Gly His Ser Thr Leu Pro Arg Thr Leu Met Ala
           -40                 -35                 -30

CCT CGG ATG ATT TCC GAG GGA GAC ATA GGA GGC ATT GCT CAA ATC ACC     335
Pro Arg Met Ile Ser Glu Gly Asp Ile Gly Gly Ile Ala Gln Ile Thr
        -25                 -20                 -15

TCC TCT CTA TTC CTG GGC AGA GGC AGT GTG GCC TCC AAT CGG CAC CTC     383
Ser Ser Leu Phe Leu Gly Arg Gly Ser Val Ala Ser Asn Arg His Leu
    -10                  -5                   1               5

CTC CAG GCT CGT GGC ATC                                             401
Leu Gln Ala Arg Gly Ile
                 10
```

(2) INFORMATION FOR SEQ ID NO: 264:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 230 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Dystrophic muscle

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..228
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 2..183
id AA022583
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(69..228)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 281..440
id AA022584
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 66..119
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq PALCLFDVDGTLT/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

```
AGGAAGTTCC GGGCCGAGTT CCTCGTGCCA ACGTGTCTTG TAAGGTGCGG CTAGAAACTG    60

GGGAC ATG GCA GCG CCT GGC CCA GCG CTC TGC CTC TTC GAC GTG GAT GGG  110
      Met Ala Ala Pro Gly Pro Ala Leu Cys Leu Phe Asp Val Asp Gly
                  -15                 -10                 -5

ACC CTC ACC GCC CCG CGG CAG AAA ATT ACC AAA GAA ATG GAT GAC TTC    158
Thr Leu Thr Ala Pro Arg Gln Lys Ile Thr Lys Glu Met Asp Asp Phe
              1               5                  10

CTA CAA AAA TTG AGG CAG AAG ATC AAA ATC GGA GTG GTA GGC GGA TCG    206
Leu Gln Lys Leu Arg Gln Lys Ile Lys Ile Gly Val Val Gly Gly Ser
     15                  20                  25

GAC TTT GAG AAA GTG CAG GAA CGG                                    230
Asp Phe Glu Lys Val Gln Glu Arg
 30                  35
```

(2) INFORMATION FOR SEQ ID NO: 265:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 224 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 101..220
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 159..278
id H97758
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 50..103
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 107..160
id H97758
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 101..185
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 166..250
id N59486
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 50..103
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 114..167
id N59486
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 50..103
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 148..201
id R09724
est

```
(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 5..54
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 2..51
                              id R09724
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 101..130
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 200..229
                              id R09724
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 101..178
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 170..247
                              id W90369
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 53..103
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 121..171
                              id W90369
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 173..218
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 91
                              region 240..285
                              id W90369
                              est

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 14..103
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 100
                              region 1..90
                              id N56221
                              est

(ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 126..182
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 3.9
                              seq ILFHGVFYAGGFA/IV
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

```
ACTGGAAGAA CTCGTCATGC TCTTTGTAGC GTGGTGCTTC TGTTGCTCAC AGGACAACTT   60

GCCTTTGATG ATTTTCAAGA GAGTTGTGCT ATGATGTGGC AAAGTATGCA GGAAGCAGGC  120

GGTCA ATG CCT CTG GGA GCA AGG ATC CTT TTC CAC GGT GTG TTC TAT GCC  170
      Met Pro Leu Gly Ala Arg Ile Leu Phe His Gly Val Phe Tyr Ala
                      -15                 -10                 -5

GGG GGC TTT GCC ATT GTG TAT TAC CTC ATT CAA AAG TTT CAT TCC AGG    218
Gly Gly Phe Ala Ile Val Tyr Tyr Leu Ile Gln Lys Phe His Ser Arg
                1               5                  10

ACA CTG                                                            224
Thr Leu
```

(2) INFORMATION FOR SEQ ID NO: 266:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 326 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..239
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..216
id HUM429E03B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 235..327
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 211..303
id HUM429E03B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 124..327
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 107..310
id T80259
est

```
(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 31..130
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 15..114
                             id T80259
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 39..283
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 95
                             region 1..245
                             id T31768
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 271..327
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                             region 234..290
                             id T31768
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 102..327
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                             region 47..272
                             id N32697
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 55..97
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                             region 1..43
                             id N32697
                             est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 65..327
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                             region 1..263
                             id N44613
                             est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 156..194
     (C) IDENTIFICATION METHOD: Von Heijne matrix
     (D) OTHER INFORMATION:  score 3.9
                             seq MLLSIGMLMLSAT/QV
```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

```
GCCTAGGTGT TGTCGTCCCT GCTAGTACTC CGGGCTGTGG GGGTCGGTGC GGATATTCAG    60

TCATGAAATC AGGGTAGGGA CTTCTCCCGC AGCGACGCGG CTGGCAAGAC TGTTTGTGTT   120

GCGGGGGCCG GACTTCAAGG TGATTTTACA ACGAG ATG CTG CTC TCC ATA GGG      173
                                       Met Leu Leu Ser Ile Gly
                                                       -10

ATG CTC ATG CTG TCA GCC ACA CAA GTC TAC ACC ATC TTG ACT GTC CAG     221
Met Leu Met Leu Ser Ala Thr Gln Val Tyr Thr Ile Leu Thr Val Gln
        -5                  1               5

CTC TTT GCA TTC TTA AAC CTA CTG CCT GTA GAA GYA GAC ATT TTA GCA     269
Leu Phe Ala Phe Leu Asn Leu Leu Pro Val Glu Xaa Asp Ile Leu Ala
 10                  15                  20                  25

TAT AAC TTT GAA AAT GCA TCT CAG ACA TTT GAT GAC CTC CCT GCA AGA     317
Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe Asp Asp Leu Pro Ala Arg
                 30                  35                  40

TTT GGT TAT                                                         326
Phe Gly Tyr
```

(2) INFORMATION FOR SEQ ID NO: 267:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 398 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 28..395
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..368
id AA150637
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 33..297
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 30..294
id H02768
est

(ix) FEATURE:

```
        (A) NAME/KEY: other
        (B) LOCATION: 181..372
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 96
                                region 148..339
                                id H70139
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 33..179
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 100
                                region 1..147
                                id H70139
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(267..394)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 90
                                region 231..358
                                id W46236
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(184..277)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 349..442
                                id W46236
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: complement(109..164)
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 464..519
                                id W46236
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 188..366
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 97
                                region 72..250
                                id N30922
                                est

    (ix) FEATURE:
        (A) NAME/KEY: other
        (B) LOCATION: 117..180
        (C) IDENTIFICATION METHOD: blastn
        (D) OTHER INFORMATION:  identity 98
                                region 1..64
                                id N30922
                                est
```

```
(ix) FEATURE:
      (A) NAME/KEY: sig_peptide
      (B) LOCATION: 111..185
      (C) IDENTIFICATION METHOD: Von Heijne matrix
      (D) OTHER INFORMATION:  score 3.9
                              seq WIAAVTIAAGTAA/IG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:

AATCGCGGAG TCGGTGCTTT AGTACGCCGC TGGCACCTTT ACTCTCGCCG GCCGCGCGAA   60

CCCGTTTGAG CTCGGTATCC TAGTGCACAC GCCTTGCAAG CGACGGCGCC ATG AGT     116
                                                       Met Ser
                                                       -25

CTG ACT TCC AGT TCC AGC GTA CGA GTT GAA TGG ATC GCA GCA GTT ACC    164
Leu Thr Ser Ser Ser Ser Val Arg Val Glu Trp Ile Ala Ala Val Thr
            -20                 -15                 -10

ATT GCT GCT GGG ACA GCT GCA ATT GGT TAT CTA GCT TAC AAA AGA TTT    212
Ile Ala Ala Gly Thr Ala Ala Ile Gly Tyr Leu Ala Tyr Lys Arg Phe
        -5                      1               5

TAT GTT AAA GAT CAT CGA AAT AAA GCT ATG ATA AAC CTT CAC ATC CAG    260
Tyr Val Lys Asp His Arg Asn Lys Ala Met Ile Asn Leu His Ile Gln
 10                 15                  20                  25

AAA GAC AAC CCC AAG ATA GTA CAT GCT TTT GAC ATG GAG GAT TTS RNA    308
Lys Asp Asn Pro Lys Ile Val His Ala Phe Asp Met Glu Asp Xaa Xaa
                30                  35                  40

GAT AAA GCT GTG TAC TGC CGT TGT TGG AGG TCC AAA AAG TTC CCA TTC    356
Asp Lys Ala Val Tyr Cys Arg Cys Trp Arg Ser Lys Lys Phe Pro Phe
            45                  50                  55

TGT GAT GGG GCT CAC ACA ARM VAT AAC GAA GAG ACT GGG CTG            398
Cys Asp Gly Ala His Thr Xaa Xaa Asn Glu Glu Thr Gly Leu
        60                      65                  70
```

(2) INFORMATION FOR SEQ ID NO: 268:

```
(i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 393 base pairs
      (B) TYPE: NUCLEIC ACID
      (C) STRANDEDNESS: DOUBLE
      (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
      (A) ORGANISM: Homo Sapiens
      (F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 55..150
      (C) IDENTIFICATION METHOD: blastn
```

(D) OTHER INFORMATION: identity 97
region 31..126
id AA094226
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 151..212
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 126..187
id AA094226
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..58
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..35
id AA094226
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 211..242
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 185..216
id AA094226
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..263
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 202..410
id R54574
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..58
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 172..206
id R54574
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..176
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 159..280
id R13710
est

(ix) FEATURE:
(A) NAME/KEY: other

(B) LOCATION: 174..235
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 279..340
id R13710
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..58
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 129..163
id R13710
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..165
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 189..299
id T78111
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 163..203
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 298..338
id T78111
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..58
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 159..193
id T78111
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 201..235
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 337..371
id T78111
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 70..252
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq YTAVSVLAGPRWA/DP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

```
AATTACGCAG AGAGAAAGTT ACGAGAAACT CGTTTTCATC TTCTTGGTTT CATCYTAAAT   60

ACCAACGTC ATG TCT GGT TCT AAT GGT TCC AAA GAA AAT TCT CAC AAT AAG  111
          Met Ser Gly Ser Asn Gly Ser Lys Glu Asn Ser His Asn Lys
              -60                 -55                 -50

GCT CGG ACG TCT CCT TAC CCA GGT TCA AAA GTT GAA CGA AGC CAG GTT    159
Ala Arg Thr Ser Pro Tyr Pro Gly Ser Lys Val Glu Arg Ser Gln Val
        -45                 -40                 -35

CCT AAT GAG AAA GTG GGC TGG CTT GTT GAG TGG CAA GAC TAT AAG CCT    207
Pro Asn Glu Lys Val Gly Trp Leu Val Glu Trp Gln Asp Tyr Lys Pro
    -30                 -25                 -20

GTG GAA TAC ACT GCA GTC TCT GTC TTG GCT GGA CCC AGG TGG GCA GAT    255
Val Glu Tyr Thr Ala Val Ser Val Leu Ala Gly Pro Arg Trp Ala Asp
-15                 -10                  -5                   1

CCT CAG ATC AGT GAV AGT VAT TTT TCT CCC AAG TTT AAC GAA AAG GAT    303
Pro Gln Ile Ser Xaa Ser Xaa Phe Ser Pro Lys Phe Asn Glu Lys Asp
              5                  10                  15

GGG CAT GTT GAG AGA NAG AGC AAG AAT GGC CTG TAT GAG ATT GAN AAT    351
Gly His Val Glu Arg Xaa Ser Lys Asn Gly Leu Tyr Glu Ile Xaa Asn
        20                  25                  30

GGA AGA CCG AGA AAT CCT GCA GAC GGA CTG GAC TGG TGG GCC            393
Gly Arg Pro Arg Asn Pro Ala Asp Gly Leu Asp Trp Trp Ala
    35                  40                  45
```

(2) INFORMATION FOR SEQ ID NO: 269:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 474 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 154..352
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 137..335
id HSC1QH021
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 154..291
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98

region 126..263
id HUML12288
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 25..111
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..87
id HUML12288
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 178..443
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..266
id R60742
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 154..303
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 133..282
id HSC07D011
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 53..147
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 35..129
id HSC07D011
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 18..49
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..32
id HSC07D011
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 154..298
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 136..280
id C04685
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 25..147

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 10..132
id C04685
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 349..438
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq LWMRWTVTSTTRA/WI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

```
AAAACCTTAG CAAGATGGCG GCTCCCTGGG CGTCCCTGCG CCTGGTCGCC CCCATGTGGA   60

ATGGGCGTAT CAGGGGCATC CATCGCCTGG GTGCGGCAGT GGCCCCAGAG GGCAATCAGA  120

AGAAGAAAAG GACAATAMTC CARKTYCCTG GMCCVAASTA TTTCTACGAT GTGGAGGCTC  180

TGAGGGATTA CTTGCTCCAA AGGGAGATGT ACAAGGTGCA TGAGAAAAAT CGATCTTACA  240

CCTGGCTGGA GAAGCAACAT GGTCCATACG GCGCAGGTGC CTTTTTCATC CTGAAGCAGG  300

GAGGCGCAGT CAAGTTTCGA GACAAGGAGT GGATCAGGCC AGATAAGT ATG GCC ATT   357
                                                     Met Ala Ile
                                                     -30

TCT CTC AGG AGT TCT GGA ATT TCT GTG AAG TGC CTG TCG AAG CTG TGG   405
Ser Leu Arg Ser Ser Gly Ile Ser Val Lys Cys Leu Ser Lys Leu Trp
        -25                 -20                 -15

ATG CGG TGG ACT GTG ACA TCA ACT ACG AGG GCC TGG ATM RNN GCN GAA   453
Met Arg Trp Thr Val Thr Ser Thr Thr Arg Ala Trp Ile Xaa Ala Glu
    -10                 -5                  1               5

CCT CCG CAG CTG GAC ATC TCG                                       474
Pro Pro Gln Leu Asp Ile Ser
                10
```

(2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 211 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 82..210

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 49..177
id AA082886
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..83
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 15..51
id AA082886
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 17..97
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq FVLGSARLGGSGS/MR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

```
AACATTAACC GGCAGG ATG TCG GAG GTG CGG CTG CCA CCG CTA CGC GCC CTG    52
                  Met Ser Glu Val Arg Leu Pro Pro Leu Arg Ala Leu
                          -25                 -20

GAC GAC TTT GTT CTG GGG TCG GCG CGT CTT GGC GGC TCC GGA TCC ATG     100
Asp Asp Phe Val Leu Gly Ser Ala Arg Leu Gly Gly Ser Gly Ser Met
-15                 -10                 -5                   1

CGA CCC GCT GCG ATG GTG YHA CCG CGT CAT CAA CAA CCT CCT CTA CTA     148
Arg Pro Ala Ala Met Val Xaa Pro Arg His Gln Gln Pro Pro Leu Leu
            5                   10                  15

CCA AAC CAA CTA CCT TCT CTG CTT CGG CAT CGG CCT CGC TCT CGC CGG     196
Pro Asn Gln Leu Pro Ser Leu Leu Arg His Arg Pro Arg Ser Arg Arg
        20                  25                  30

GTA CGT ACG GCC ACG                                                 211
Val Arg Thr Ala Thr
    35
```

(2) INFORMATION FOR SEQ ID NO: 271:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 262 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 184..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 177..257
id W93162
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 113..183
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 107..177
id W93162
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 184..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 174..254
id W67415
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 62..112
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 52..102
id W67415
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 141..183
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 132..174
id W67415
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 184..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 174..254
id N44655
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 113..183
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 104..174
id N44655

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..183
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 41..170
id HSBA7H051
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 184..240
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 170..226
id HSBA7H051
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..52
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 2.,40
id HSBA7H051
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 77..183
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 63..169
id R37538
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 184..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 169..249
id R37538
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 15..53
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..39
id R37538
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 206..250
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8

seq LVSATAWLEECWW/SE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:

```
AGGCGGCGAA GATGGCGGAG AACAGCGGTC GGGCGGGCAA GAGCAGCGGG AGCGNCGCGN    60

GGAAGGGGGC GGTGTCCGCA GAGCAGGTGA TTGCTGGCTT CAACCGCCTT CGGCAGGAAC   120

AGCGAGGCCT GGCATCCAAA GCAGCTGAGT TGGAGATGGA GTTGAATGAG CACAGCCTAG   180

TGAATCGATA CACTGAAGGA GGTAG ATG AAA CTC GTA AGT GCT ACC GCA TGG     232
                            Met Lys Leu Val Ser Ala Thr Ala Trp
                            -15                 -10

TTG GAG GAR TGC TGG TGG AGC GAA CTG TCA                             262
Leu Glu Glu Cys Trp Trp Ser Glu Leu Ser
    -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 272:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 422 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 142..382
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 120..360
id HUML1108
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 29..139
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 17..127
id HUML1108
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 37..395
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 37..395
id AA156844
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 32..395
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 137..500
id HSU51712
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 237..395
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 112..270
id T70871
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 133..235
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 9..111
id T70871
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 77..185
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 105..213
id H48308
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 177..286
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 206..315
id H48308
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 284..317
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 314..347
id H48308
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 309..410
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LYVPLLAVCCLHS/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

```
AAGCTTCCAA ACCCAGGGCT TGCGCTTGCC TTTGCCTCTT CCACCGGGCA GGGACCATGT    60

CGGCGGAGAC CGCGAGCGGC CCCACAGAGG ACCAGGTGGA AATCCTGGAG TACAACTTCA   120

ACAAGGTCGA CAAGCACCCG GATTCCACCA CGCTGTGCCT CATCGCGGCC GAGGCAGGCC   180

TTTCCGAGGA GGAGACCCAG AAATGGTTTA AGCAGCGCCT GGCAAAGTGG CGGCGCTCAG   240

AAGGCCTGCC CTCAGAGTGC AGATCCGTCA CAGACTAAGG AGATGGCAGG CATTGACAGC   300

TTCACTCC ATG AAG GCC ATC TCT GTT TCT CTC CTC CGC TTA ACC AAG CTG     350
         Met Lys Ala Ile Ser Val Ser Leu Leu Arg Leu Thr Lys Leu
                         -30                 -25

TTG TGG TTT TTC AGC ATA GTG TTG TAT GTT CCA TTG CTA GCT GTC TGC      398
Leu Trp Phe Phe Ser Ile Val Leu Tyr Val Pro Leu Leu Ala Val Cys
-20                 -15                 -10                  -5

TGT TTA CAC AGT GTT GTA TTT TTT                                      422
Cys Leu His Ser Val Val Phe Phe
                 1
```

(2) INFORMATION FOR SEQ ID NO: 273:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 513 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Thyroid

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 195..421
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 179..405
id AA010986
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 20..109
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 2..91
id AA010986
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 108..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 91..188
id AA010986
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 443..505
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 429..491
id AA010986
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 417..449
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 402..434
id AA010986
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 19..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 1..187
id W96112
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 316..494
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 300..478
id W96112
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 195..336
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 178..319
id W96112
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 69..513
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..445
id W44481
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 2..193
id AA129812
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 195..300
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 184..289
id AA129812
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 349..405
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 340..396
id AA129812
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 301..352
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 291..342
id AA129812
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 405..448
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 397..440
id AA129812
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..290
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..289
id W40172
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 342..439
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 343..440

id W40172
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 285..342
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 285..342
id W40172
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 85..438
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LMIALTVVGCIFM/VI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

```
ACTCCAACGC TGGGTGACAT TGAGCTCACC AGCGCCACCG TCCCCGGCGA AGTTCTGCGC    60

TGGTCGGCGG AGTAGCAAGT GGCC ATG GGG AGC CTC AGC GGT CTG CGC CTG      111
                           Met Gly Ser Leu Ser Gly Leu Arg Leu
                                       -115                -110

GCA GCA GGA AGC TGT TTT AGG TTA TGT GAA AGA GAT GTT TCC TGN TCT     159
Ala Ala Gly Ser Cys Phe Arg Leu Cys Glu Arg Asp Val Ser Xaa Ser
            -105                -100                -95

CTA AGG CTT ACC AGA AGC TCT GAT TTA AAG AGA ATA AAT GGA TTT TGC     207
Leu Arg Leu Thr Arg Ser Ser Asp Leu Lys Arg Ile Asn Gly Phe Cys
        -90                 -85                 -80

ACA AAA CCA CAG GAA AGT CCC GGA GCT CCA TCC CGC ACT TAC AAC AGA     255
Thr Lys Pro Gln Glu Ser Pro Gly Ala Pro Ser Arg Thr Tyr Asn Arg
    -75                 -70                 -65

GTG CCT TTA CAC AAA CCT ACG GAT TGG CAG AAA AAG ATC CTC ATA TGG     303
Val Pro Leu His Lys Pro Thr Asp Trp Gln Lys Lys Ile Leu Ile Trp
    -60                 -55                 -50

TCA GGT CGC TTC AAA AAG GAA ANB NAA ATC CCA GAG ACT GTC TCG TTG     351
Ser Gly Arg Phe Lys Lys Glu Xaa Xaa Ile Pro Glu Thr Val Ser Leu
-45                 -40                 -35                 -30

GAG ATG CTT GAN STT GCA AAG AAC AAG ATG CGA GTG AAG ATC AGC TAT     399
Glu Met Leu Xaa Xaa Ala Lys Asn Lys Met Arg Val Lys Ile Ser Tyr
                -25                 -20                 -15

CTA ATG ATT GCC CTG ACG GTG GTA GGA TGC ATC TTC ATG GTT ATT GAG     447
Leu Met Ile Ala Leu Thr Val Val Gly Cys Ile Phe Met Val Ile Glu
            -10                 -5                  1

GGC AAG AAG GCT GCC CAA AGA CAC GAG ACT TTA ACA AGC TTG MAC TTA     495
Gly Lys Lys Ala Ala Gln Arg His Glu Thr Leu Thr Ser Leu Xaa Leu
    5                   10                  15

GAA AAG AAA GCT CGT CTG                                             513
```

Glu Lys Lys Ala Arg Leu
20 25

(2) INFORMATION FOR SEQ ID NO: 274:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 412 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 198..407
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 190..399
id AA001815
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 41..147
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 35..141
id AA001815
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 146..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 139..198
id AA001815
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 198..400
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 139..341
id N42162
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 60..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 95
region 2..147

id N42162
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 198..354
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 144..300
id N24414
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 62..147
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 10..95
id N24414
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 146..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 93..152
id N24414
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 198..414
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 122..338
id W76137
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 75..147
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..73
id W76137
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 146..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 71..130
id W76137
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 198..360
(C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 96
region 121..283
id H03817
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 77..147
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 2..72
id H03817
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 346..402
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 270..326
id H03817
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 146..205
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 91
region 70..129
id H03817
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 59..358
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq LASSFLFTMGGLG/FI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 274:

```
ACTGTTTNNG GGAGGCGCGT GGGGCTTGAG GCCGAGAACG GCCCTTGCTG CCACCAAC      58

ATG GAG ACT TTG TAC CGT GTC CCG TTC TTA GTG CTC GAA TGT GCC AAC     106
Met Glu Thr Leu Tyr Arg Val Pro Phe Leu Val Leu Glu Cys Pro Asn
-100                -95                 -90                 -85

CTG AAG CTG AAG AAG CCG CCC TGG TTG CAC ATG CCG TCG GCC ATG ACT     154
Leu Lys Leu Lys Lys Pro Pro Trp Leu His Met Pro Ser Ala Met Thr
                -80                 -75                 -70

GTG TAT GCT CTG GTG GTG GTG TCT TAC TTC CTC ATC ACC GGA GGA ATA     202
Val Tyr Ala Leu Val Val Val Ser Tyr Phe Leu Ile Thr Gly Gly Ile
            -65                 -60                 -55

ATT TAT GAT GTT ATT GTT GAA CCT CCA AGT GTC GGT TCT ATG ACT GAT     250
Ile Tyr Asp Val Ile Val Glu Pro Pro Ser Val Gly Ser Met Thr Asp
        -50                 -45                 -40

GAA CAT GGG CAT CAG AGG CCA GTA GCT TTC TTG GCC TAC AGA GTA AAT     298
```

```
Glu His Gly His Gln Arg Pro Val Ala Phe Leu Ala Tyr Arg Val Asn
    -35                 -30                 -25

GGA CAA TAT ATT ATG GAA GGA CTT GCA TCC AGC TTC CTA TTT ACA ATG      346
Gly Gln Tyr Ile Met Glu Gly Leu Ala Ser Ser Phe Leu Phe Thr Met
-20                 -15                 -10                  -5

GGA GGT TTA GGT TTC ATA ATC CTG GAC GGA TCG RNT GCA CCA AAT ATC      394
Gly Gly Leu Gly Phe Ile Ile Leu Asp Gly Ser Xaa Ala Pro Asn Ile
                 1               5                  10

CCA AAA CTC AAT AGA TTC                                              412
Pro Lys Leu Asn Arg Phe
        15
```

(2) INFORMATION FOR SEQ ID NO: 275:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 243 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 136..238
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 80..182
id C05215
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 73..111
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq MLVLRSGLTKALA/SR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 275:

```
CACTCGGGAA GACTTCAGAG AAGTCTCACA AAGGACTCGG CTGGCTGCTT TTCTCAGTGC     60

CGAAGCCGCG CC ATG CTC GTT CTC AGA AGC GGC CTG ACC AAG GCG CTT GCC    111
              Met Leu Val Leu Arg Ser Gly Leu Thr Lys Ala Leu Ala
                          -10                 -5

TCA CGG ACG CTC GCG CVT CAG AKA AWT TTT GCT CAT CGA GCT GAA GTT      159
Ser Arg Thr Leu Ala Xaa Gln Xaa Xaa Phe Ala His Arg Ala Glu Val
 1               5                  10                  15

CGG AAA GCC TTA GCC AAC TGT AAG GAA TGG CAA GAA CAA TCT ATC ATT      207
```

```
Arg Lys Ala Leu Ala Asn Cys Lys Glu Trp Gln Glu Gln Ser Ile Ile
            20                  25                  30

CCA AAT TTG GCT CGC ATT GAT AAA CAA GAG ACC AGG                 243
Pro Asn Leu Ala Arg Ile Asp Lys Gln Glu Thr Arg
        35                  40
```

```
(2) INFORMATION FOR SEQ ID NO: 276:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 245 base pairs
          (B) TYPE: NUCLEIC ACID
          (C) STRANDEDNESS: DOUBLE
          (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Homo Sapiens
          (F) TISSUE TYPE: Thyroid

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 112..241
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 98
                                  region 77..206
                                  id R87832
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 37..113
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 92
                                  region 1..77
                                  id R87832
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 112..241
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 98
                                  region 65..194
                                  id HUM427G10B
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
          (B) LOCATION: 49..113
          (C) IDENTIFICATION METHOD: blastn
          (D) OTHER INFORMATION:  identity 98
                                  region 1..65
                                  id HUM427G10B
                                  est

    (ix) FEATURE:
          (A) NAME/KEY: other
```

(B) LOCATION: 112..241
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 52..181
id R52722
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 62..113
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 1..52
id R52722
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 111..241
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 79..209
id W41484
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 30..137
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq NIESLAWTGGTLG/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 276:

```
GAGTTTCCTG CGAGCTCGGC TTCCTCAAC ATG GCT GCG CCC TTG TCA GTG GAG       53
                                Met Ala Ala Pro Leu Ser Val Glu
                                    -35                 -30

GTG GAG TTC GGA GGT GGT GCG GAS TCC TGT TTG ACG GTA TTA AGA AAC      101
Val Glu Phe Gly Gly Gly Ala Xaa Ser Cys Leu Thr Val Leu Arg Asn
            -25                 -20                 -15

ATC GAG TCA CTT GCC TGG ACA GGA GGA ACC CTG GGA CAT CCG GAA CCT      149
Ile Glu Ser Leu Ala Trp Thr Gly Gly Thr Leu Gly His Pro Glu Pro
        -10                 -5                   1

GCT CAT CTG GAT CAA GAA GAA TTT GCT AAA AGA GCG GCC ASA GTT GTT      197
Ala His Leu Asp Gln Glu Glu Phe Ala Lys Arg Ala Ala Xaa Val Val
  5                 10                  15                  20

CAT CCA GGG AGA CAG CGT GCG GCC AGG AAT TCT GGT GCT GAC TAC AGG      245
His Pro Gly Arg Gln Arg Ala Ala Arg Asn Ser Gly Ala Asp Tyr Arg
                25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 277:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 401 base pairs

(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 22..403
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..382
id AA127626
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 64..349
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 44..329
id W39584
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 349..403
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 330..384
id W39584
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 24..60
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 3..39
id W39584
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(47..403)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 68..424
id N32838
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(56..403)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 67..414

id AA121528
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 164..378
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 130..344
id AA082078
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 36..165
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..130
id AA082078
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 198..392
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq FVGGLPVIFWSWA/GL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 277:

```
ACTTAGTCGT GTGTACATCA TTGGGAATGG AGGGAAATAA ATGACTGGAT GGTCGCTGCT    60

TTTTAAGTTT CAAATTGACA TTCCAGACAA GCGGTGCCTG AGCCTGTGCC TGTCTTCAGA   120

TCTTCACAGC ACAGTTCCTG GGAAGGTGGA GCCACCAGCC TCTCCTTGAA TAACTGGGAG   180

ATGAAACAGG AAGCTCT ATG ACA CAC TTG ATC GAA TAT GAC AGA CAC CGA      230
                   Met Thr His Leu Ile Glu Tyr Asp Arg His Arg
                   -65                 -60                 -55

AAA TCA CGA CTC AGC CCC CTC CAG CAC CTC TAC CTG TTG CCC GCC GAT     278
Lys Ser Arg Leu Ser Pro Leu Gln His Leu Tyr Leu Leu Pro Ala Asp
                -50                 -45                 -40

CAC AGC CGG AAT GCA GCT GAA AGA TTC CCT GGG GCC TGG TTC CAA CCG     326
His Ser Arg Asn Ala Ala Glu Arg Phe Pro Gly Ala Trp Phe Gln Pro
            -35                 -30                 -25

CCC ACT GTG GAC TCT GAG GCC TCT GCA TTT GTG GGT GGT CTG CCT GTG     374
Pro Thr Val Asp Ser Glu Ala Ser Ala Phe Val Gly Gly Leu Pro Val
        -20                 -15                 -10

ATA TTT TGG TCA TGG GCT GGT CTG GTC                                 401
Ile Phe Trp Ser Trp Ala Gly Leu Val
     -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 278:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 335 base pairs
    (B) TYPE: NUCLEIC ACID
    (C) STRANDEDNESS: DOUBLE
    (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
    (A) ORGANISM: Homo Sapiens
    (F) TISSUE TYPE: Uterus

(ix) FEATURE:
    (A) NAME/KEY: other
    (B) LOCATION: 70..337
    (C) IDENTIFICATION METHOD: blastn
    (D) OTHER INFORMATION: identity 97
        region 1..268
        id HSC2SG081
        est

(ix) FEATURE:
    (A) NAME/KEY: other
    (B) LOCATION: 71..251
    (C) IDENTIFICATION METHOD: blastn
    (D) OTHER INFORMATION: identity 98
        region 1..181
        id R13964
        est

(ix) FEATURE:
    (A) NAME/KEY: other
    (B) LOCATION: 256..334
    (C) IDENTIFICATION METHOD: blastn
    (D) OTHER INFORMATION: identity 100
        region 184..262
        id R13964
        est

(ix) FEATURE:
    (A) NAME/KEY: other
    (B) LOCATION: 26..255
    (C) IDENTIFICATION METHOD: blastn
    (D) OTHER INFORMATION: identity 94
        region 1..230
        id HUML13589
        est

(ix) FEATURE:
    (A) NAME/KEY: other
    (B) LOCATION: 116..251
    (C) IDENTIFICATION METHOD: blastn
    (D) OTHER INFORMATION: identity 96
        region 1..136
        id H05572
        est

(ix) FEATURE:
    (A) NAME/KEY: other
    (B) LOCATION: 256..337
    (C) IDENTIFICATION METHOD: blastn

(D) OTHER INFORMATION: identity 100
region 139..220
id H05572
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 24..89
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq WARKLLSVPWLLC/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 278:

```
AACAGTTACG CGCCGCACGG ATC ATG GCC GCA GCC GCT CTG GGG CAG ATC TGG    53
                          Met Ala Ala Ala Ala Leu Gly Gln Ile Trp
                                  -20                 -15

GCA CGA AAG CTT CTC TCT GTC CCT TGG CTT CTG TGT GGT CCC AGA AGA     101
Ala Arg Lys Leu Leu Ser Val Pro Trp Leu Leu Cys Gly Pro Arg Arg
    -10                     -5                   1

TAT GCC TCC TCC AGT TTC AAG GCT GCA GAC CTG CAG CTG GAA ATG ACA     149
Tyr Ala Ser Ser Ser Phe Lys Ala Ala Asp Leu Gln Leu Glu Met Thr
  5                  10                  15                  20

CAG AAG CCT CAT AAG AAG CCT GGC CCC GGC GAG CCC CTG GTG TTT GGG     197
Gln Lys Pro His Lys Lys Pro Gly Pro Gly Glu Pro Leu Val Phe Gly
                 25                  30                  35

AAG ACA TTT ACC GAC CAC ATG CTG ATG GTG GAA TGG AAT GAC AAG GGC     245
Lys Thr Phe Thr Asp His Met Leu Met Val Glu Trp Asn Asp Lys Gly
             40                  45                  50

TGG GGC CAG CCC CGA ATC CAG CCC TTC CAG AAC CTC ACG CTG CAC CCA     293
Trp Gly Gln Pro Arg Ile Gln Pro Phe Gln Asn Leu Thr Leu His Pro
         55                  60                  65

GCC TCC TCC AGC CTC CAC TAC TCC CTG CAG CTG TTT GAG GGC             335
Ala Ser Ser Ser Leu His Tyr Ser Leu Gln Leu Phe Glu Gly
     70                  75                  80
```

(2) INFORMATION FOR SEQ ID NO: 279:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 344 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: other

```
      (B) LOCATION: 57..176
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 50..169
                              id AA126817
                              est

  (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 219..344
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 94
                              region 213..338
                              id AA126817
                              est

  (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 10..344
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 2..336
                              id W79731
                              est

  (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 19..344
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 1..326
                              id H21245
                              est

  (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 31..302
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 96
                              region 34..305
                              id H11314
                              est

  (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 302..344
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 97
                              region 306..348
                              id H11314
                              est

  (ix) FEATURE:
      (A) NAME/KEY: other
      (B) LOCATION: 41..202
      (C) IDENTIFICATION METHOD: blastn
      (D) OTHER INFORMATION:  identity 98
                              region 22..183
                              id W19587
                              est
```

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 201..284
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 183..266
id W19587
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 283..344
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 266..327
id W19587
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 48..161
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq CPLLLLVFTTNNG/RH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 279:

```
AAGGGGTCGG AGGTCAGGGC GAGCGTCTCG CAGGCCGTAG GAGGAAG ATG GCG GTG      56
                                                    Met Ala Val

GAG TCG CGC GTT ACC CAG GAG GAA ATT AAG AAG GAG CCA GAG AAA CCG     104
Glu Ser Arg Val Thr Gln Glu Glu Ile Lys Lys Glu Pro Glu Lys Pro
-35                 -30                 -25                 -20

ATC GAC CGC GAG AAG ACA TGC CCA CTG TTG CTA CTG GTC TTC ACC ACC     152
Ile Asp Arg Glu Lys Thr Cys Pro Leu Leu Leu Leu Val Phe Thr Thr
                -15                 -10                 -5

AAT AAC GGC CGC CAC CAC CGA ATG GAC GAG TTC TCC CGG GGA AAT GTA     200
Asn Asn Gly Arg His His Arg Met Asp Glu Phe Ser Arg Gly Asn Val
              1               5                  10

CCG TCC AGC GAG TTG CAG ATC TAC ACT TGG ATG GAT GCA ACT TTG AAA     248
Pro Ser Ser Glu Leu Gln Ile Tyr Thr Trp Met Asp Ala Thr Leu Lys
     15                  20                  25

GAA CTG ACA AGC TTA GTA AAA GAA GTC TAC CCA GAA GCT AGA WAG AAG     296
Glu Leu Thr Ser Leu Val Lys Glu Val Tyr Pro Glu Ala Arg Xaa Lys
 30                  35                  40                  45

GGC ACT CAC TTC AAT TTT GCA VTC GTT TTT ACA GAT GTT AAA AGA CCT     344
Gly Thr His Phe Asn Phe Ala Xaa Val Phe Thr Asp Val Lys Arg Pro
                 50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO: 280:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 401 base pairs

(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 111..377
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 72..338
id W79829
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 370..401
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 332..363
id W79829
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 111..377
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 14..280
id H62624
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 370..401
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 274..305
id H62624
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 111..377
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 24..290
id H81957
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 111..376
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92

```
                                    region 59..324
                                    id W82998
                                    est

    (ix) FEATURE:
         (A) NAME/KEY: other
         (B) LOCATION: 111..376
         (C) IDENTIFICATION METHOD: blastn
         (D) OTHER INFORMATION:  identity 92
                                 region 64..329
                                 id AA023811
                                 est

    (ix) FEATURE:
         (A) NAME/KEY: sig_peptide
         (B) LOCATION: 240..305
         (C) IDENTIFICATION METHOD: Von Heijne matrix
         (D) OTHER INFORMATION:  score 3.7
                                 seq AVLDCAFYDPTHA/WS

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 280:

ACTAGCCTGC GAGTGTTCTG AGGGAAGCAA GGAGGCGGCG GCGGCCGCAG CGAGTGGCGA    60

GTAGTGGAAA CGTTGCTTCT GAGGGGTGTC CAAGATGASC GGTTCKAMCG GAGKTCAAGC   120

TGAACCAGCC ACCCGAGGAT GGCATCTCCT CCGTGAAGTT CAGCCCCAAC ACCTCCCAGT   180

TCCTGCTTGT CTCCTCCTGG GACACGTCCG TGCGTCTCTA CGATGTGCCG GCCAACTCC    239

ATG CGG CTC AAG TAC CAG CAC ACC GGC GCC GTC CTG GAC TGC GCC TTC     287
Met Arg Leu Lys Tyr Gln His Thr Gly Ala Val Leu Asp Cys Ala Phe
        -20                 -15                 -10

TAC GAT CCA ACG CAT GCC TGG AGT GGA GGA CTA GAT CAT CAA TTG AAA     335
Tyr Asp Pro Thr His Ala Trp Ser Gly Gly Leu Asp His Gln Leu Lys
     -5                  1               5                  10

ATG CAT GAT TTG AAC ACT GAT CAA GAA AAT CTT GTT GGG ACC ATG ATG     383
Met His Asp Leu Asn Thr Asp Gln Glu Asn Leu Val Gly Thr Met Met
                 15                  20                  25

CCC CTA TCA GAT GTG TTG                                             401
Pro Leu Ser Asp Val Leu
             30


(2) INFORMATION FOR SEQ ID NO: 281:

     (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 275 base pairs
          (B) TYPE: NUCLEIC ACID
          (C) STRANDEDNESS: DOUBLE
          (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: CDNA

    (vi) ORIGINAL SOURCE:
          (A) ORGANISM: Homo Sapiens
```

(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 87..272
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 72..257
id T60345
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..89
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 33..75
id T60345
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..47
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..34
id T60345
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 87..272
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 75..260
id T46853
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 12..89
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..78
id T46853
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 87..207
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 39..159
id R57601
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 193..272
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92

region 144..223
id R57601
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 48..89
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 1..42
id R57601
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 84..195
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 90
region 55..166
id W71083
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 12..269
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq WAVVLADTAVTSG/RG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 281:

```
ATAGGCGCAA G ATG GCG CTG CTT TTT GCA CGT TCT TTG CGC TTG TGC CGC      50
             Met Ala Leu Leu Phe Ala Arg Ser Leu Arg Leu Cys Arg
                 -85                 -80                 -75

TGG GGA GCC AAA CGA TTG GGA GTT GCC TCC ACA GAG GCC CAG AGA GGC      98
Trp Gly Ala Lys Arg Leu Gly Val Ala Ser Thr Glu Ala Gln Arg Gly
            -70                 -65                 -60

GTC AGT TTC AAA CTG GMA GAA AAA ACC GCC CAC AGC AGC CTG GCA CTC     146
Val Ser Phe Lys Leu Xaa Glu Lys Thr Ala His Ser Ser Leu Ala Leu
        -55                 -50                 -45

TTC AGA GAT GAT ACG GGT GTC AAA TAT GGC TTG GTG GGA TTG GAG CCC     194
Phe Arg Asp Asp Thr Gly Val Lys Tyr Gly Leu Val Gly Leu Glu Pro
    -40                 -35                 -30

ACC AAG GTG GCC TTG AAT GTG GAG CGC TTC CGG GAG TGG GCA GTG GTG     242
Thr Lys Val Ala Leu Asn Val Glu Arg Phe Arg Glu Trp Ala Val Val
-25                 -20                 -15                 -10

CTG GCA GAC ACA GCG GTC ACC AGT GGC AGA GGG                         275
Leu Ala Asp Thr Ala Val Thr Ser Gly Arg Gly
                -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 282:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 397 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE.
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 77..280
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq ILLGNYCVAVADA/KK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 282:

```
ATTCCCCCTT GGGCGGTGGT GGAGGTGGTA ACCGTGATAG TAGCAGCTCC GGCGGCAGCA    60

ACAGCGACTA CGAGGG ATG GCG GCG GCT GCA GCA GGA ACT SNA ACA TCC CAG   112
                  Met Ala Ala Ala Ala Ala Gly Thr Xaa Thr Ser Gln
                                  -65                 -60

AGG TTT TTC CAG AGC TTC TCG GAT GCC CTA ATC GAC GAG GAC CCC CAG      160
Arg Phe Phe Gln Ser Phe Ser Asp Ala Leu Ile Asp Glu Asp Pro Gln
    -55                 -50                 -45

GCG GCG TTA GAG GAG CTG ACT AAG GCT TTG GAA CAG AAA CCA GAT GAT      208
Ala Ala Leu Glu Glu Leu Thr Lys Ala Leu Glu Gln Lys Pro Asp Asp
-40                 -35                 -30                 -25

GCA CAG TAT TAT TGT CAA AGA GCT TAT TGT CAC ATT CTT CTT GGG AAT      256
Ala Gln Tyr Tyr Cys Gln Arg Ala Tyr Cys His Ile Leu Leu Gly Asn
                -20                 -15                 -10

TAC TGT GTT GCT GTT GCT GAT GCA AAG AAG TCT CTA GAA CTC AAT CCA      304
Tyr Cys Val Ala Val Ala Asp Ala Lys Lys Ser Leu Glu Leu Asn Pro
            -5                  1                   5

AAT AAT TCC ACT GCT ATG CTG AGA AAA GGA ATA TGT GAA TAC CAT GAA      352
Asn Asn Ser Thr Ala Met Leu Arg Lys Gly Ile Cys Glu Tyr His Glu
    10                  15                  20

AAA AAC TAT GCT GCT GCC CTA GAA ACT TTT TAC AGA AGG ACG GGG          397
Lys Asn Tyr Ala Ala Ala Leu Glu Thr Phe Tyr Arg Arg Thr Gly
25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 283:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 381 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 78..379
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 79..380
id H17763
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..53
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 4..55
id H17763
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 96..377
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 82..363
id H16532
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..53
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 7..58
id H16532
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 79..370
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 46..337
id R52491
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 66..248
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 65..247
id R21494
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 2..53
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..52
id R21494
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 266..305
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 268..307
id R21494
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 129..321
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 71..263
id AA084554
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 315..379
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 256..320
id AA084554
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 139..318
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq WFYIGSSLNGTRG/KR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:

```
AGTGGCCCGG ATGTTCGGTG CAGCTGCCAG ATCCGCTGAT CTAGTGCTTC TGGAAAAAAA    60

CCTTCAGGCG GCCCATGGCT GTCGATATTC AACCAGCATG CCTTGGACTT TATTSYGGGA   120

AGACCCTATT ATTTAAAA ATG GCT CAA CTG AAA TAT ATG GAG AAT GTG GGG     171
                    Met Ala Gln Leu Lys Tyr Met Glu Asn Val Gly
                    -60                 -55                 -50

TAT GCC CAA GAG GAC AGA GAA CGA ATG CAC AGA AAT ATT GTC AGC CTT     219
Tyr Ala Gln Glu Asp Arg Glu Arg Met His Arg Asn Ile Val Ser Leu
                -45                 -40                 -35

GCA CAG AAT CTC CTG AAC TTT ATG ATT GGC TCT ATC TTG GAT TTA TGG     267
Ala Gln Asn Leu Leu Asn Phe Met Ile Gly Ser Ile Leu Asp Leu Trp
            -30                 -25                 -20
```

```
CAA TGC TTC CTC TGG TTT TAC ATT GGT TCT TCA TTG AAT GGT ACT CGG      315
Gln Cys Phe Leu Trp Phe Tyr Ile Gly Ser Ser Leu Asn Gly Thr Arg
        -15                 -10                 -5

GGA AAA AGA GTT CCA GCG CAC TTT TCC AAC ACA TCA CTG CAT TAT TTG      363
Gly Lys Arg Val Pro Ala His Phe Ser Asn Thr Ser Leu His Tyr Leu
    1               5                   10                  15

AAT GCA GCA TGG CCG CGG                                              381
Asn Ala Ala Trp Pro Arg
                20
```

(2) INFORMATION FOR SEQ ID NO: 284:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 293 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 3..294
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..292
id HUM524F05B
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 44..172
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 48..176
id H81799
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 167..276
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 170..279
id H81799
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 14..48
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97

```
                              region 16..50
                              id H81799
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 48..172
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 99
                              region 57..181
                              id T84779
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 167..226
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 93
                              region 175..234
                              id T84779
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 1..45
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 97
                              region 7..51
                              id T84779
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 167..294
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 94
                              region 101..228
                              id W81213
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 66..172
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                              region 1..107
                              id W81213
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 8..172
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 98
                              region 1..165
                              id AA090080
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 167..210
```

(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 159..202
id AA090080
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 174..266
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq WSPLSTRSGGTHA/CS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:

```
AAAAACAATA GGACGGAAAC GCCGAGGAAC CCGGCTGAGG CGGCAGAGCA TCCTGGCCAG   60

AACAAGCCAA GGAGCCAAGA CGAGAGGGAC ACACGGACAA ACAACAGACA GAAGACGTAC  120

TGGCCGCTGG ACTCCKCTGC CTCCCCCATC TCCCCGCCAT CTGCGCCCGG AGG ATG     176
                                                           Met

AGC CCA GCC TTC AGG GCC ATG GAT GTG GAG CCC CGC GCC AAA GGS TCC     224
Ser Pro Ala Phe Arg Ala Met Asp Val Glu Pro Arg Ala Lys Gly Ser
-30                 -25                 -20                 -15

TTC TGG AGC CCT TTG TCC ACC AGG TCG GGG GGC ACT CAT GCG TGC TCC     272
Phe Trp Ser Pro Leu Ser Thr Arg Ser Gly Gly Thr His Ala Cys Ser
                -10                  -5                   1

GCT TCA ATG AGA CAA CCC TGG                                         293
Ala Ser Met Arg Gln Pro Trp
         5
```

(2) INFORMATION FOR SEQ ID NO: 285:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 347 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 26..326
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 42..342
id R71425
est

(ix) FEATURE:

(A) NAME/KEY: other
(B) LOCATION: 19..345
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 11..337
id AA133412
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(114..345)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 172..403
id AA156940
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(71..114)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 402..445
id AA156940
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: complement(26..76)
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 439..489
id AA156940
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 172..345
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 186..359
id W07240
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 72..171
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 83..182
id W07240
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 21..76
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 92
region 33..88
id W07240
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 39..345
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..307
id R81019
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 18..179
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq SILAQVLDQSARA/RL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:

```
AGCGCTGACG CCGAGCC ATG GCG GAC GAG GAG CTT GAG GCG CTG AGG AGA        50
                   Met Ala Asp Glu Glu Leu Glu Ala Leu Arg Arg
                                   -50                 -45

CAG AGG CTG GCC GAG CTG CAG GCC AAA CAC GGG GAT CCT GGT GAT GCG       98
Gln Arg Leu Ala Glu Leu Gln Ala Lys His Gly Asp Pro Gly Asp Ala
            -40                 -35                 -30

GCC CAA CAG GAA GCA AAG CAC AGG GAA GCA GAA ATG AGA AAC AGT ATC      146
Ala Gln Gln Glu Ala Lys His Arg Glu Ala Glu Met Arg Asn Ser Ile
        -25                 -20                 -15

TTA GCC CAA GTT CTG GAT CAG TCG GCC CGG GCC AGG TTA AGT AAC TTA      194
Leu Ala Gln Val Leu Asp Gln Ser Ala Arg Ala Arg Leu Ser Asn Leu
    -10                 -5                   1               5

GCA CTT GTA AAG CCT GAA AAA ACT AAA GCA GTA GAG AAT TAC CTT ATA      242
Ala Leu Val Lys Pro Glu Lys Thr Lys Ala Val Glu Asn Tyr Leu Ile
                10                  15                  20

CAG ATG GCA AGA TAT GGA CAA CTA AGT GAG AAG GTA TCA GAA CAA GGT      290
Gln Met Ala Arg Tyr Gly Gln Leu Ser Glu Lys Val Ser Glu Gln Gly
            25                  30                  35

TTA ATA GAR ATC CTT AAA AAA GTA AGC CAA CAA ACA GAA AAG AHN ACA      338
Leu Ile Glu Ile Leu Lys Lys Val Ser Gln Gln Thr Glu Lys Xaa Thr
        40                  45                  50

ACA GTG AGG                                                          347
Thr Val Arg
    55
```

(2) INFORMATION FOR SEQ ID NO: 286:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 414 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 186..382
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 156..352
id AA082259
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 61..146
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 34..119
id AA082259
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 29..61
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 1..33
id AA082259
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 194..331
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 169..306
id H80945
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..146
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 93
region 30..122
id H80945
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 157..345
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq GLVCAGLADMARP/AE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:

```
AACAGCGGGC AGGGAAAGCC GCGGGAAGGG TACTCCAGGC GAGAGGCGGA CGCGAGTCGT    60

CGTGGCAGGA AAAGTGACTA GCTCCCCTTC GTTGTCAGCC AGGGACGAGA ACACAGCCAC   120

GCTCCCAMCC GGCTGCCHAA GRWTCCCTSG GCGGCG ATG TCG GCC GCC GGT GCC     174
                                        Met Ser Ala Ala Gly Ala
                                                    -60

CGA GGC CTG CGG GCC ACC TAC CAC CGG CTC CTC GAT AAA GTG GAG CTG     222
Arg Gly Leu Arg Ala Thr Tyr His Arg Leu Leu Asp Lys Val Glu Leu
        -55                 -50                 -45

ATG CTG CCC GAG AAA TTG AGG CCG TTG TAC AAC CAT CCA GCA GGT CCC     270
Met Leu Pro Glu Lys Leu Arg Pro Leu Tyr Asn His Pro Ala Gly Pro
    -40                 -35                 -30

AGA ACA GTT TTC TTC TGG GCT CCA ATT ATG AAA TGG GGG TTG GTG TGT     318
Arg Thr Val Phe Phe Trp Ala Pro Ile Met Lys Trp Gly Leu Val Cys
-25                 -20                 -15                 -10

GCT GGA TTG GCT GAT ATG GCC AGA CCT GCA GAA AAA CTT AGC ACA GCT     366
Ala Gly Leu Ala Asp Met Ala Arg Pro Ala Glu Lys Leu Ser Thr Ala
                -5                  1                   5

CAA TCT GVK GTT TTG ATG GCT ACA GGG TTT ATT TGG TCA AGA TAC TCG     414
Gln Ser Xaa Val Leu Met Ala Thr Gly Phe Ile Trp Ser Arg Tyr Ser
        10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 287:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 478 base pairs
  (B) TYPE: NUCLEIC ACID
  (C) STRANDEDNESS: DOUBLE
  (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
  (A) ORGANISM: Homo Sapiens
  (F) TISSUE TYPE: Muscle

(ix) FEATURE:
  (A) NAME/KEY: other
  (B) LOCATION: 196..391
  (C) IDENTIFICATION METHOD: blastn
  (D) OTHER INFORMATION: identity 95
    region 185..380
    id W07314
    est

(ix) FEATURE:
  (A) NAME/KEY: other
  (B) LOCATION: 58..204
  (C) IDENTIFICATION METHOD: blastn
  (D) OTHER INFORMATION: identity 98
    region 46..192
    id W07314

est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 409..478
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 401..470
id W07314
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 34..412
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 12..390
id W07582
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 45..393
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 31..379
id W73850
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 16..52
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 1..37
id W73850
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 73..447
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..375
id AA112776
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 63..388
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..326
id H72671
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 98..355
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5

seq TGXLNMTLQRASA/AP

(xi) SEQUENCE DESCRIPTION: SEQ.ID NO: 287:

```
AACTTGTCAG CCCTTGTCTG AGGCGGAGGC AGCCCCGCGC CGCGCCGGAC CCGAGCATAT   60

TTCATTTTCT GTCATTGGAC TTTGAGCCAT TAGAACC ATG AGC AAC TAC AGT GTG   115
                                         Met Ser Asn Tyr Ser Val
                                             -85

TCA CTG GTT GGC CCA GCT CCT TGG GGT TTC CGG CTG CAG GGC GGT AAG   163
Ser Leu Val Gly Pro Ala Pro Trp Gly Phe Arg Leu Gln Gly Gly Lys
-80                 -75                 -70                 -65

GAT TTC AAC ATG CCT CTG ACA ATC TCT AGT CTA AAA GAT GGC GGC AAG   211
Asp Phe Asn Met Pro Leu Thr Ile Ser Ser Leu Lys Asp Gly Gly Lys
                -60                 -55                 -50

GCA GCC CAG GCA AAT GTA AGA ATA GGC GAT GTG GTT CTC AGC ATT GAT   259
Ala Ala Gln Ala Asn Val Arg Ile Gly Asp Val Val Leu Ser Ile Asp
            -45                 -40                 -35

GGA ATA AAT GCA CAA GGA ATG ACT CAT CTT GAA GCC CAG AAT AAG ATT   307
Gly Ile Asn Ala Gln Gly Met Thr His Leu Glu Ala Gln Asn Lys Ile
        -30                 -25                 -20

AAG GGT TGT ACA GGA NYT TTG AAT ATG ACT CTG CAA AGA GCA TCT GCT   355
Lys Gly Cys Thr Gly Xaa Leu Asn Met Thr Leu Gln Arg Ala Ser Ala
    -15                 -10                 -5

GCA CCC AAG CCT GAG CCG GTT CCT GTT CAA AAG CCC ACA GTC ACC AGC   403
Ala Pro Lys Pro Glu Pro Val Pro Val Gln Lys Pro Thr Val Thr Ser
1               5                   10                  15

GTG TGT TCC GAG ACT TCT CAG GAG CTA GCA GAG GGA CAG AGA AGA GGA   451
Val Cys Ser Glu Thr Ser Gln Glu Leu Ala Glu Gly Gln Arg Arg Gly
            20                  25                  30

TCC CAG GGT GAC AGT AAA CAG CAA AAT                               478
Ser Gln Gly Asp Ser Lys Gln Gln Asn
        35                  40
```

(2) INFORMATION FOR SEQ ID NO: 288:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: other

(B) LOCATION: 4..333
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..330
id N35568
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 26..297
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..272
id R35915
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 295..338
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 271..314
id R35915
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 44..255
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..212
id W31312
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 251..355
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 209..313
id W31312
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 21..328
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 13..320
id HSC1MA011
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 62..339
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 1..278
id R61491
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 245..298
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq LLGLELSEAEAIG/AD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 288:

```
ATTCGTTTAC AGTTCGGCAC GTAGGACGGA GGGTAGTGCG TCTAGAGACA CATATTCCCA    60

ACGGATTTGA CGATGGTGTT CGGTCTTGAA TGGAAATGTA GTCTTAGGCC AGTCTTAGGT   120

TTTTGAACAG GATAGTAGGT ATCCGGAGTC GATTGAGGGC CAGAGCAGGC ACTGGGGTTC   180

GGATCCTGGG CAAAGTTTCC CACATTGAGG GTCTCGAGGA CGCCTAGATC TCTTTCCCAG   240

GGCC ATG GCG AAC CCG AAG CTG CTG GGA CTG GAG CTA AGC GAG GCG GAG    289
     Met Ala Asn Pro Lys Leu Leu Gly Leu Glu Leu Ser Glu Ala Glu
                 -15                 -10                 -5

GCG ATC GGT GCT GAT TCG GCG CGA TTT GAG GAG CTG CTG CTG CAG GCC     337
Ala Ile Gly Ala Asp Ser Ala Arg Phe Glu Glu Leu Leu Leu Gln Ala
              1               5                  10

TCG AAG GAG CTC CAG CAA                                             355
Ser Lys Glu Leu Gln Gln
    15
```

(2) INFORMATION FOR SEQ ID NO: 289:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 401 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 113..201
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 90..178
id W21198
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 23..74
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98

```
                              region 2..53
                              id W21198
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: 71..111
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 100
                              region 49..89
                              id W21198
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(114..201)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 94
                              region 271..358
                              id AA061731
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(114..201)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 94
                              region 271..358
                              id AA061768
                              est

(ix) FEATURE:
     (A) NAME/KEY: other
     (B) LOCATION: complement(125..201)
     (C) IDENTIFICATION METHOD: blastn
     (D) OTHER INFORMATION:  identity 96
                              region 269..345
                              id AA058174
                              est

(ix) FEATURE:
     (A) NAME/KEY: sig_peptide
     (B) LOCATION: 204..323
     (C) IDENTIFICATION METHOD: Von Heijne matrix
     (D) OTHER INFORMATION:  score 3.5
                              seq ALLCTLLLHFQNI/RR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:

AAAGGTGTCT GGATCGGAGG GAGGTTCGGG TGGGCATCGG GCGGCTGGAA GAGCTCGACT   60

CGTCCCGCTG GGAAAGCGCG AGTCTGAGTG GAACCCTGGA CGACTTGCAG AGCGGCTGGC  120

GCAGTCATGG CGGACTACTG GAAGTCACAG CCAAAGAAAT TCTGTGATTA CTGCAAGTGC  180

TGGATAGCAG ACAATAGGCC TGT ATG ATA ATT CCG CTG TTA GAG ATT CTA ATA  233
                          Met Ile Ile Pro Leu Leu Glu Ile Leu Ile
                          -40                 -35

ATA ATT GTG TTG AAT GAA GTG CTC CTT TTT GAT GTA AAC TCA GTT TAC    281
```

```
Ile Ile Val Leu Asn Glu Val Leu Leu Phe Asp Val Asn Ser Val Tyr
-30                 -25                 -20                 -15

AAA GCA CTT TTA TGT ACA TTG CTC TTG CAT TTT CAA AAC ATC AGA AGA      329
Lys Ala Leu Leu Cys Thr Leu Leu Leu His Phe Gln Asn Ile Arg Arg
                -10                 -5                   1

TTT CTG TCT TCT CAG TCC CCT ATG AAA GCT GTG AGC CTT CTA THT TTT      377
Phe Leu Ser Ser Gln Ser Pro Met Lys Ala Val Ser Leu Leu Xaa Phe
            5                   10                  15

CAT CAA CCT GAC TTT GAT TAT ATA                                      401
His Gln Pro Asp Phe Asp Tyr Ile
     20                 25
```

(2) INFORMATION FOR SEQ ID NO: 290:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 385 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 50..382
(C) IDENTIFICATION METHOD: fasta
(D) OTHER INFORMATION: identity 97
region 4..337
id HUMGPCRB
vrt

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 292..345
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 1..54
id T29782
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 345..382
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 94
region 55..92
id T29782
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 80..235

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq LVFIIGLVGNLLA/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290:

```
AACTTCAGTT TGGACAACTA CTCACAGCTA CTACACAGAG ACCCGAACGA GTCACTGATA    60

TACACCTGGA CCACCACCA ATG GAT ATA CAA ATG GCA AAC AAT TTT ACT CCG   112
                     Met Asp Ile Gln Met Ala Asn Asn Phe Thr Pro
                             -50                 -45

CCC TCT GCA ACT CCT CAG GGA AAT GAC TGT GAC CTC TAT GCA CAT CAC   160
Pro Ser Ala Thr Pro Gln Gly Asn Asp Cys Asp Leu Tyr Ala His His
    -40                 -35                 -30

AGC ACG GCC AGG ATA GTA ATG CCT CTG CAT TAC AGC CTC GTC TTC ATC   208
Ser Thr Ala Arg Ile Val Met Pro Leu His Tyr Ser Leu Val Phe Ile
-25                 -20                 -15                 -10

ATT GGG CTC GTG GGA AAC TTA CTA GCC TTG GTC GTC ATT GTT CAA AAC   256
Ile Gly Leu Val Gly Asn Leu Leu Ala Leu Val Val Ile Val Gln Asn
                -5                   1                   5

AGG AAA AAA ATC AAC TCT ACC ACC CTC TAT TCA ACA AAT TTG GTT ATT   304
Arg Lys Lys Ile Asn Ser Thr Thr Leu Tyr Ser Thr Asn Leu Val Ile
        10                  15                  20

TCT GAT ATA CTT TTT ACC ACC GCT TTG CCT ACA CGA ATA GCT ACT ATG   352
Ser Asp Ile Leu Phe Thr Thr Ala Leu Pro Thr Arg Ile Ala Thr Met
    25                  30                  35

CMA TGG GCY TTG ACT GGA GAA TCG GAG ATG TGG                       385
Xaa Trp Ala Leu Thr Gly Glu Ser Glu Met Trp
40                  45                  50
```

(2) INFORMATION FOR SEQ ID NO: 291:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 461 base pairs
(B) TYPE: NUCLEIC ACID
(C) STRANDEDNESS: DOUBLE
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: CDNA

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Pancreas

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..462
(C) IDENTIFICATION METHOD: fasta
(D) OTHER INFORMATION: identity 99
region 1..408
id HUMORF06
vrt

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 47..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..218
id W77946
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 263..412
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 216..365
id W77946
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 412..462
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 366..416
id W77946
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..462
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..409
id C16991
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 263..462
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 212..411
id N28784
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 102..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 97
region 52..214
id N28784
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 50..107
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..58

id N28784
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 54..356
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 99
region 1..303
id C17735
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 357..462
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 303..408
id C17735
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 102..264
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 96
region 47..209
id AA057588
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 263..406
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 98
region 207..350
id AA057588
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 406..462
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 351..407
id AA057588
est

(ix) FEATURE:
(A) NAME/KEY: other
(B) LOCATION: 55..107
(C) IDENTIFICATION METHOD: blastn
(D) OTHER INFORMATION: identity 100
region 1..53
id AA057588
est

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: 357..443
(C) IDENTIFICATION METHOD: Von Heijne matrix

```
        (D) OTHER INFORMATION:  score 7
                                seq SMIGIGSLPSCWA/CW

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291:

AGTTCGTTTA TTCCTCCGCG CGCTGGGACA GGCTGCTTCT TCGCCAGAAC CAACCGGTTG   60

CTTGCTGTCC CAGCGGCGCC CCCTCATCAC CGTCGCCATG CCCGGAGGTC TGCTTCTCGG  120

GGACGTGGCT CCCAACTTTG AGGCCAATAC CACCGTCGGC CGCATCCGTT TCCACGACTT  180

TCTGGGAKAC TCATGGGGCA TTCTCTTCTC CCACCCTCGG GACTTTACCC CAGTGTGCAC  240

CACAGAGCTT GGCAGAGCTG CAAAGCTGGC ACCAGAATTT GCCAAGAGGA ATGTTAAGTT  300

GWTTGCCCTT TCAATAGACA GTGTTGAGGA CCATCTTGCC TGGAGCAAGG ATATCA ATG  359
                                                              Met

CTT ACA ATT GTG AAG AGC CCA CAG AAA AGT TAC CTT TTC CCA TCA TCG    407
Leu Thr Ile Val Lys Ser Pro Gln Lys Ser Tyr Leu Phe Pro Ser Ser
            -25                 -20                 -15

ATG ATA GGA ATC GGG AGC TTG CCA TCC TGT TGG GCA TGC TGG ATC CAG    455
Met Ile Gly Ile Gly Ser Leu Pro Ser Cys Trp Ala Cys Trp Ile Gln
        -10                 -5                   1

CAG AGA                                                            461
Gln Arg
  5
```

(2) INFORMATION FOR SEQ ID NO: 292:

```
    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 69 amino acids
        (B) TYPE: AMINO ACID
        (D) TOPOLOGY: LINEAR

    (ii) MOLECULE TYPE: PROTEIN

    (vi) ORIGINAL SOURCE:
        (A) ORGANISM: Homo Sapiens
        (F) TISSUE TYPE: Liver

    (ix) FEATURE:
        (A) NAME/KEY: sig_peptide
        (B) LOCATION: -35..-1
        (C) IDENTIFICATION METHOD: Von Heijne matrix
        (D) OTHER INFORMATION:  score 15
                                seq LFLLLLLAASAWG/VT

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:

Met Ser Ser Trp Ser Arg Gln Arg Pro Lys Ser Pro Gly Gly Ile Gln
-35                 -30                 -25                 -20

Pro His Val Ser Arg Thr Leu Phe Leu Leu Leu Leu Leu Ala Ala Ser
                -15                 -10                 -5
```

Ala Trp Gly Val Thr Leu Ser Pro Lys Asp Cys Gln Val Phe Arg Ser
1 5 10

Asp His Gly Ser Ser Ile Ser Cys Gln Pro Pro Ala Glu Ile Pro Gly
15 20 25

Tyr Leu Pro Ala Thr
30

(2) INFORMATION FOR SEQ ID NO: 293:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 13.2
seq LLLXAVLLSLASA/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 293:

Met Arg Val Arg Ile Gly Leu Thr Leu Leu Leu Xaa Ala Val Leu Leu
-20 -15 -10

Ser Leu Ala Ser Ala Ser Ser Asp Glu Glu Gly Ser Gln Asp Glu Ser
-5 1 5 10

Leu Asp Ser Lys Thr Thr Leu Thr Ser Asp Glu Ser Val Lys Asp His
15 20 25

Thr Thr Ala Gly Arg Val Val Ala Gly Gln Ile Phe Leu Asp Ser Glu
30 35 40

Glu Ser Glu Leu Glu Xaa Ser Ile Gln Glu Glu Glu Asp Ser Leu Lys
45 50 55

Ser Gln Glu Gly Glu Ser Val Thr Glu Asp Ile Ser Phe Leu Glu Ser
60 65 70 75

(2) INFORMATION FOR SEQ ID NO: 294:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 13.1
seq CVLLLLLLLLTRS/SE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294:

```
Met Phe Ser His Leu Pro Phe Asp Cys Val Leu Leu Leu Leu Leu Leu
    -20                 -15                 -10

Leu Leu Thr Arg Ser Ser Glu Val Glu Xaa Xaa Ala Glu Val Gly Gln
 -5                   1               5                  10

Asn Ala Tyr Leu Pro Cys Phe Tyr Thr Pro Ala Ala Pro Gly Asn Leu
             15                  20                  25

Val Pro Val Cys Trp Gly Lys Gly Ala Cys Pro Val Phe Glu Cys Gly
         30                  35                  40

Asn Val Val Leu Arg Thr Asp Glu Arg Asp Val Asn Tyr Trp Thr Ser
     45                  50                  55

Arg Tyr Trp Leu Asn Gly Asp Phe Arg Lys Gly Asp Val Ser Leu Thr
 60                  65                  70                  75

Ile Glu Asn Val Thr Leu Ala Asp Ser Gly Ile Tyr Cys Cys Arg Ile
                 80                  85                  90

Gln Ile Pro Gly Ile Met Asn Asp Glu Lys Phe Asn Leu
             95                 100
```

(2) INFORMATION FOR SEQ ID NO: 295:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 11.6
seq LLFLFLAVDEAWA/GM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295:

```
Met Gly Pro Val Arg Leu Gly Ile Leu Leu Phe Leu Phe Leu Ala Val
    -20                 -15                 -10

Asp Glu Ala Trp Ala Gly Met Leu Lys Glu Glu Gly Arg
 -5                     1               5
```

(2) INFORMATION FOR SEQ ID NO: 296:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.7
seq SLLLAVALGLATA/VS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:

```
Met Lys Ser Leu Ser Leu Leu Leu Ala Val Ala Leu Gly Leu Ala Thr
        -15                 -10                 -5

Ala Val Ser Ala Gly Pro Ala Val Ile Glu Cys Trp Phe Val Glu Asp
     1                   5                  10                  15

Ala Ser Gly Lys Gly Leu Ala Lys Arg Pro Gly Ala Leu Leu Leu Arg
                20                  25                  30

Gln Gly Pro Gly Glu Pro Pro Pro Arg Pro Asp Leu Asp Pro Glu Leu
            35                  40                  45

Tyr Leu Ser Val His Asp Pro Ala Gly Ala Leu Gln Ala Arg
        50                  55                  60
```

(2) INFORMATION FOR SEQ ID NO: 297:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 105 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.6
seq LLTLXLLGGPTWA/GK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:

```
Met Leu Leu Leu Leu Thr Leu Xaa Leu Leu Gly Gly Pro Thr Trp Ala
    -15                 -10                 -5

Gly Lys Met Tyr Gly Pro Gly Gly Gly Lys Tyr Phe Ser Thr Thr Glu
  1               5                  10                  15

Asp Tyr Asp His Glu Ile Thr Gly Leu Arg Val Ser Val Gly Leu Leu
             20                  25                  30

Leu Val Lys Ser Val Gln Val Lys Leu Gly Asp Ser Trp Asp Val Lys
         35                  40                  45

Leu Gly Ala Leu Xaa Gly Asn Thr Gln Glu Val Xaa Xaa Gln Pro Gly
     50                  55                  60

Glu Tyr Ile Thr Lys Val Phe Val Ala Phe Gln Ala Phe Leu Arg Gly
 65                  70                  75                  80

Met Val Met Tyr Thr Ser Lys Asp Arg
                 85
```

(2) INFORMATION FOR SEQ ID NO: 298:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 77 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -46..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.4
seq LIILIXIWIWCLG/SQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:

```
Met Lys Ile Gly Ile Leu Leu Ser Leu Leu Asn Ser Val Ile Ser Gln
    -45                 -40                 -35

Thr Leu Met Ser Cys Asn Trp Lys Gln Gln Met Arg Arg Met Lys Thr
-30                 -25                 -20                 -15

Ile Leu Ile Ile Leu Ile Xaa Ile Trp Ile Trp Cys Leu Gly Ser Gln
                -10                 -5                   1

Thr Phe Gly Thr Ser Thr Thr Lys Ser Val Gln Leu Lys Ile Leu Arg
        5                   10                  15

Gln Asn Leu Ser His Phe Leu Gln Pro Pro Gln Val Ile
    20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO: 299:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -30..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.4
seq LPFLLSLFPGALP/VQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 299:

```
Met Lys Ala Ser Ser Gly Arg Cys Gly Leu Val Arg Trp Leu Gln Val
-30                 -25                 -20                 -15

Leu Leu Pro Phe Leu Leu Ser Leu Phe Pro Gly Ala Leu Pro Val Gln
                -10                 -5                   1

Ile Arg Tyr Ser Ile Pro Glu Glu Leu Ala Lys Asn Ser Val Val Gly
        5                   10                  15

Asn Leu Ala Lys Asp Leu Gly Leu Ser Val Arg Asp Leu Pro Ala Arg
    20                  25                  30

Lys Leu Arg Val Ser Ala Glu Lys Glu Tyr Phe Thr Val Asn Pro Glu
35                  40                  45                  50

Ser Gly Asp Leu Leu Val Ser Asp Arg Ile Asp Arg Asp Val
                55                  60
```

(2) INFORMATION FOR SEQ ID NO: 300:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -33..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.1
seq IIFLCHLLRGLHA/XT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:

```
Met Ile Val Asp Cys Val Ser Ser His Leu Lys Lys Thr Gly Asp Gly
            -30                 -25                 -20

Ala Lys Thr Phe Ile Ile Phe Leu Cys His Leu Leu Arg Gly Leu His
        -15                 -10                 -5

Ala Xaa Thr Asp Arg Glu Lys Asp Pro Leu Met Cys Glu Asn Ile Gln
    1               5                   10                  15

Thr His Gly Arg Leu Pro
                20
```

(2) INFORMATION FOR SEQ ID NO: 301:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 109 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Ovary

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -104..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.1
seq LTSLSWLLXASCS/KP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:

```
Met Ala Lys Ala Leu Leu Phe Pro Ser Gly Arg Ser Val Arg Val Leu
```

```
                -100                -95                 -90

Tyr Gly Ala Val Asn Lys Glu Arg Gln Xaa Glu Ser Val Leu Asn Arg
            -85                 -80                 -75

Ala Cys Pro Pro Lys Ala Asn Ser Lys Glu Arg Arg Gly Arg Ala Val
        -70                 -65                 -60

Leu Gly Ala Glu Leu Thr Gln Trp Ser Ser Pro Thr Thr Ala Gly Ser
    -55                 -50                 -45

Cys Cys Ser Ser Cys Thr Leu Cys Ala Arg Ser Ser Ser Xaa Val Ile
-40                 -35                 -30                     -25

Ala Pro Ser Pro Leu Val Pro Phe Thr Ser Gly Leu Thr Ser Leu Ser
                -20                 -15                 -10

Trp Leu Leu Xaa Ala Ser Cys Ser Lys Pro Xaa Lys Gly
            -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 302:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -73..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8
seq LATKLLSLSGVFA/VH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:

```
Met Ala Ala Ser Glu Ala Ala Val Val Ser Ser Pro Ser Leu Lys Thr
            -70                 -65                 -60

Asp Thr Ser Pro Val Leu Glu Thr Ala Gly Thr Val Ala Ala Met Ala
        -55                 -50                 -45

Ala Thr Pro Ser Ala Arg Ala Ala Ala Ala Val Val Ala Ala Ala Ala
    -40                 -35                 -30

Arg Thr Gly Ser Glu Ala Arg Val Ser Lys Ala Ala Leu Ala Thr Lys
-25                 -20                 -15                     -10

Leu Leu Ser Leu Ser Gly Val Phe Ala Val His Lys Pro Lys Gly Pro
                -5                  1               5

Thr Ser Ala Glu Leu Leu Asn Arg Leu Lys Glu Lys Leu Leu Ala Glu
```

```
                10                      15                      20

Ala Gly Met Pro Ser Pro Glu Trp Thr Xaa Arg Lys Lys Gln Thr Xaa
    25                      30                      35

Glu Asn Trp Ala Trp Arg Asp Ser Arg Gln Arg Xaa Arg Gly Val Leu
40                      45                      50                      55

Val Val Gly Ile Gly Ala
                60
```

(2) INFORMATION FOR SEQ ID NO: 303:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.8
seq VLWLISFFTFTDG/HG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 303:

```
Met Lys Val Gly Val Leu Trp Leu Ile Ser Phe Phe Thr Phe Thr Asp
        -15                     -10                     -5

Gly His Gly Gly Phe Leu Gly Val Ser Trp Cys Tyr Val Ser Tyr Leu
    1                   5                   10                  15

Phe Ser Thr Asn Ser Pro Leu Ser Phe Arg Arg Met
                20                  25
```

(2) INFORMATION FOR SEQ ID NO: 304:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq WIFLAAILKGVQC/EV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:

```
Met Glu Phe Gly Leu Ser Trp Ile Phe Leu Ala Ala Ile Leu Lys Gly
                -15                 -10                 -5

Val Gln Cys Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Lys
             1               5                  10

Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Asp Phe
     15                  20                  25

Thr Asp Ala Trp Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu
 30                  35                  40                  45

Glu Trp Val Ala Asn Ile Xaa Ser Thr Ala Ser Gly Gly Thr Arg Gly
                 50                  55                  60

Tyr Ala Ala Pro Val Lys Asp Arg Phe Ile Ile Ser Arg Asp Asp Ser
             65                  70                  75

Arg Asn Thr Leu His Leu Gln Met Asn Gly Leu Lys Xaa Met Thr Gln
         80                  85                  90

Ala Ile Tyr Tyr Cys Ala Thr
     95                 100
```

(2) INFORMATION FOR SEQ ID NO: 305:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 150 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq LWRLLLWAGTAFQ/VX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:

```
Met Ala Glu Pro Gly His Ser His His Leu Ser Ala Arg Val Arg Gly
    -35                 -30                 -25
```

Arg Thr Glu Arg Arg Ile Pro Arg Leu Trp Arg Leu Leu Leu Trp Ala
    -20   -15   -10

Gly Thr Ala Phe Gln Val Xaa Gln Gly Xaa Xaa Pro Glu Leu Xaa Ala
-5   1   5   10

Cys Lys Glu Ser Glu Tyr His Tyr Glu Tyr Thr Ala Cys Asp Ser Thr
            15   20   25

Gly Ser Arg Trp Arg Val Ala Val Pro His Thr Xaa Gly Leu Cys Thr
        30   35   40

Ser Leu Pro Asp Pro Val Lys Gly Thr Glu Cys Xaa Xaa Ser Cys Asn
    45   50   55

Ala Gly Glu Phe Leu Asp Met Lys Asp Gln Ser Cys Xaa Pro Cys Ala
60   65   70   75

Glu Gly Arg Tyr Ser Leu Gly Thr Gly Ile Arg Phe Asp Glu Trp Asp
                80   85   90

Glu Leu Pro His Gly Phe Ala Ala Ser Gln Pro Thr Trp Ser Trp Met
            95   100   105

Thr Val Leu Leu Ser His
        110

(2) INFORMATION FOR SEQ ID NO: 306:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 105 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq QACLLGLFALILS/GK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:

Met Thr Ala Asp Pro Arg Lys Gly Arg Met Gly Leu Gln Ala Cys Leu
-25   -20   -15   -10

Leu Gly Leu Phe Ala Leu Ile Leu Ser Gly Lys Cys Ser Xaa Ser Pro
                -5   1   5

Glu Pro Asp Gln Arg Arg Thr Leu Pro Pro Gly Trp Val Ser Leu Gly
        10   15   20

Arg Ala Asp Pro Glu Glu Glu Leu Ser Leu Thr Phe Ala Leu Arg Gln
25 30 35

Gln Asn Val Glu Arg Leu Ser Glu Leu Val Gln Ala Val Ser Asp Pro
40 45 50 55

Ser Ser Pro Gln Tyr Gly Lys Tyr Leu Thr Leu Glu Asn Val Ala Asp
60 65 70

Leu Val Arg Pro Ser Pro Leu Thr Pro
75 80

(2) INFORMATION FOR SEQ ID NO: 307:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 87 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq LCFLLLAVAMSFF/GS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:

Met Leu Val Asp Gly Pro Ser Glu Arg Pro Ala Leu Cys Phe Leu Leu
-20 -15 -10

Leu Ala Val Ala Met Ser Phe Phe Gly Ser Ala Leu Ser Ile Asp Glu
-5 1 5

Thr Arg Ala His Leu Leu Leu Lys Xaa Lys Met Met Arg Leu Gly Gly
10 15 20

Arg Leu Val Leu Asn Thr Lys Glu Glu Leu Ala Asn Glu Arg Leu Met
25 30 35 40

Thr Leu Xaa Ile Ala Glu Met Lys Glu Ala Met Arg Thr Leu Ile Phe
45 50 55

Pro Pro Ser Met His Phe Phe
60

(2) INFORMATION FOR SEQ ID NO: 308:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 128 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.8
seq LVLVLVVAVTVRA/AL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:

```
Met Ala Ala Pro Leu Val Leu Val Leu Val Val Ala Val Thr Val Arg
        -15                 -10                     -5

Ala Ala Leu Phe Arg Ser Ser Leu Ala Glu Phe Ile Ser Glu Arg Val
     1                   5                  10                  15

Glu Val Val Ser Pro Leu Ser Ser Trp Lys Arg Val Val Glu Gly Leu
                 20                  25                  30

Ser Leu Leu Asp Leu Gly Val Ser Pro Tyr Ser Gly Ala Val Phe His
             35                  40                  45

Glu Thr Pro Leu Ile Ile Tyr Leu Phe His Phe Leu Ile Asp Tyr Ala
         50                  55                  60

Glu Leu Val Phe Met Ile Thr Asp Ala Leu Thr Ala Ile Ala Leu Tyr
     65                  70                  75

Phe Ala Ile Gln Asp Phe Asn Lys Val Val Phe Lys Lys Gln Lys Leu
 80                  85                  90                  95

Leu Leu Glu Leu Asp Gln Tyr Ala Pro Asp Val Ala Glu Leu Ile Arg
                100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 309:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: -102..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.7
seq LXMTLMLPFKILS/DS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:

```
Met Thr Ala Ala Ile Arg Arg Gln Arg Glu Leu Ser Ile Leu Pro Lys
        -100                -95                 -90

Val Thr Leu Glu Ala Met Asn Thr Thr Val Met Gln Gly Phe Asn Arg
    -85                 -80                 -75

Ser Glu Arg Cys Pro Arg Asp Thr Arg Ile Val Gln Leu Val Phe Pro
-70                 -65                 -60                 -55

Ala Leu Tyr Thr Val Val Phe Leu Thr Gly Ile Leu Leu Asn Thr Leu
                -50                 -45                 -40

Ala Leu Trp Val Phe Val His Ile Pro Ser Ser Ser Thr Phe Ile Ile
            -35                 -30                 -25

Tyr Leu Lys Asn Thr Leu Val Ala Asp Leu Xaa Met Thr Leu Met Leu
        -20                 -15                 -10

Pro Phe Lys Ile Leu Ser Asp Ser His Leu Ala Pro Trp Gln Leu Arg
    -5                  1               5                   10

Ala Phe Val Cys Arg Phe Ser Ser Val Ile Phe Tyr Glu Thr Met Tyr
                15                  20                  25

Val Gly Glu Gly
            30
```

(2) INFORMATION FOR SEQ ID NO: 310:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -46..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq SIGVLTLSHLISG/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:

```
Met Ser Ser Val Leu Ala Ala Ser His Pro Leu Val Leu Ser Ser Asn
```

-45 -40 -35

Ala Gly Thr Pro Gly Ile Ser Glu Lys Asp Asn Arg Asp Pro Ala Gly
-30 -25 -20 -15

Ser Ser Ile Gly Val Leu Thr Leu Ser His Leu Ile Ser Gly Leu Arg
-10 -5 1

Thr Leu Tyr Thr Leu Leu His Phe Pro Leu Arg
5 10

(2) INFORMATION FOR SEQ ID NO: 311:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 105 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Thyroid

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -50..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq LIILGLVLFMVYG/NV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311:

Met Gly Leu Ala Met Glu His Gly Gly Ser Tyr Ala Arg Ala Gly Gly
-50 -45 -40 -35

Ser Ser Arg Gly Cys Trp Tyr Tyr Leu Arg Tyr Phe Phe Leu Phe Val
-30 -25 -20

Ser Leu Ile Gln Phe Leu Ile Ile Leu Gly Leu Val Leu Phe Met Val
-15 -10 -5

Tyr Gly Asn Val His Val Ser Thr Glu Ser Asn Leu Gln Ala Thr Glu
1 5 10

Arg Arg Ala Glu Gly Leu Tyr Xaa Gln Leu Leu Gly Leu Thr Ala Ser
15 20 25 30

Gln Ser Asn Leu Thr Lys Glu Leu Asn Phe Thr Thr Arg Ala Lys Asp
35 40 45

Ala Ile Met Gln Met Trp Leu Asn Ala
50 55

(2) INFORMATION FOR SEQ ID NO: 312:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -64..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq SCLVSGWGLLANG/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

```
Met Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro
                -60                 -55                 -50

Leu Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser
            -45                 -40                 -35

Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr
        -30                 -25                 -20

Ala Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly
    -15                 -10                  -5

Gln Arg
  1
```

(2) INFORMATION FOR SEQ ID NO: 313:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 142 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -47..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq VICCVLFLLFILG/YI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:

```
Met Gly Gly Lys Gln Arg Asp Glu Asp Asp Glu Ala Tyr Gly Lys Pro
        -45                 -40                 -35

Val Lys Tyr Asp Pro Ser Phe Arg Gly Pro Ile Lys Asn Arg Ser Cys
    -30                 -25                 -20

Thr Asp Val Ile Cys Cys Val Leu Phe Leu Leu Phe Ile Leu Gly Tyr
-15                 -10                 -5                      1

Ile Val Val Gly Ile Val Ala Trp Leu Tyr Gly Asp Pro Arg Gln Val
            5                   10                  15

Leu Tyr Pro Arg Asn Ser Thr Gly Ala Tyr Cys Gly Met Gly Glu Asn
        20                  25                  30

Lys Asp Lys Pro Tyr Leu Leu Tyr Phe Asn Ile Phe Ser Cys Ile Leu
    35                  40                  45

Ser Ser Asn Ile Ile Ser Val Ala Glu Asn Gly Leu Gln Cys Pro Thr
50                  55                  60                  65

Pro Gln Val Cys Val Ser Ser Cys Pro Glu Asp Pro Trp Thr Xaa Xaa
                70                  75                  80

Lys Thr Ser Ser His Arg Leu Leu Gly Lys Ser Ser Ile Gln
            85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO: 314:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq VLLFLAWVCFLFY/AG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314:

```
Met Gln Lys Ala Ser Val Leu Leu Phe Leu Ala Trp Val Cys Phe Leu
            -15                 -10                 -5

Phe Tyr Ala Gly Ile Ala Leu Phe Thr Ser Gly Phe Leu Leu Thr Arg
        1                   5                   10

Leu Glu Leu Thr Asn His Ser Ser Cys Gln Glu Pro Pro Gly Pro Gly
15                  20                  25                  30
```

Ser Leu Pro Trp Gly Ser Gln Gly Lys Pro Gly Ala Cys Trp Met Ala
35 40 45

Ser Arg Phe Ser Arg Val Val Leu Val Leu Ile Asp Ala Leu Arg Phe
50 55 60

Asp Phe Ala Gln Pro Gln His Ser His Val Pro Arg Glu Pro Pro Val
65 70 75

Ser Leu Pro Phe Leu Gly Lys Leu Ser Ser Leu Gln Arg Ile Leu Glu
80 85 90

Ile Gln Pro His His Ala Arg Leu
95 100

(2) INFORMATION FOR SEQ ID NO: 315:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -81..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq CWMMLLGSXGSFL/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:

Met Ser Pro Val Leu His Phe Tyr Val Arg Pro Ser Gly His Glu Gly
-80 -75 -70

Ala Ala Ser Gly His Thr Arg Arg Lys Leu Gln Gly Lys Leu Pro Glu
-65 -60 -55 -50

Leu Gln Gly Val Glu Thr Glu Leu Cys Tyr Asn Val Asn Trp Thr Ala
-45 -40 -35

Glu Ala Leu Pro Ser Ala Glu Glu Thr Lys Lys Leu Met Trp Leu Phe
-30 -25 -20

Gly Cys Pro Tyr Cys Trp Met Met Leu Leu Gly Ser Xaa Gly Ser Phe
-15 -10 -5

Leu Ala Pro Met Thr Cys Xaa Trp Arg Ser
1 5

(2) INFORMATION FOR SEQ ID NO: 316:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq ILRLLGSLSNAYS/PR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:

```
Met Asp Val Thr Pro Arg Glu Ser Leu Ser Ile Leu Val Val Ala Gly
    -35                 -30                 -25

Ser Gly Gly His Thr Thr Glu Ile Leu Arg Leu Leu Gly Ser Leu Ser
-20                 -15                 -10                  -5

Asn Ala Tyr Ser Pro Arg His Tyr Val Ile Ala Asp Thr Asp Glu Met
                  1               5                  10

Ser Ala Asn Lys Ile Asn Ser Phe Glu Leu Asp Arg Ala Asp Arg Asp
         15                  20                  25

Pro Ser Asn Met Tyr Thr Lys Tyr Tyr Ile His Arg Asn Gly
     30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 317:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LLRVLNLPHNSIG/CV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:

```
Met Met Gly Val Ala Lys Leu Thr Leu Leu Arg Val Leu Asn Leu Pro
    -20                 -15                 -10

His Asn Ser Ile Gly Cys Val Glu Gly Leu Lys Glu Leu Val His Leu
 -5                  1               5                  10

Glu Trp Leu Asn Leu Ala Gly Asn Asn Leu Lys Ala Met Glu Gln Xaa
            15                  20                  25

Asn Ser Cys Thr Ala Leu Gln His Leu Asp
        30                  35
```

(2) INFORMATION FOR SEQ ID NO: 318:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq ILRLLGSLSNAYS/PR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:

```
Met Asp Val Thr Pro Arg Glu Ser Leu Ser Ile Leu Val Val Ala Gly
    -35                 -30                 -25

Ser Gly Gly His Thr Thr Glu Ile Leu Arg Leu Leu Gly Ser Leu Ser
-20                 -15                 -10                  -5

Asn Ala Tyr Ser Pro Arg His Tyr Val Ile Ala Asp Thr Asp Glu Met
                 1               5                  10

Ser Ala Asn Lys Ile Asn Ser Phe Glu Leu Asp Arg Ala Asp Arg Asp
        15                  20                  25

Arg
```

(2) INFORMATION FOR SEQ ID NO: 319:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Colon

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq MVLLTMIARVADG/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:

```
Met Val Leu Leu Thr Met Ile Ala Arg Val Ala Asp Gly Leu Pro Leu
            -10                  -5                   1

Ala Ala Ser Met Gln Glu Glu Val Arg Thr Ala Pro Arg Ala Leu
        5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 320:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -47..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq GCGMFTFLSSVXA/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:

```
Met Val Pro Val Glu Asn Thr Glu Gly Pro Ser Leu Leu Asn Gln Lys
        -45                 -40                 -35

Gly Thr Ala Val Glu Thr Glu Gly Xaa Gly Ser Arg His Pro Pro Trp
    -30                 -25                 -20

Ala Arg Gly Cys Gly Met Phe Thr Phe Leu Ser Ser Val Xaa Ala Ala
-15                 -10                  -5                   1

Val Ser Gly Leu Leu Val Gly Tyr Glu Leu Gly Ile Ile Ser Gly Ala
            5                   10                  15

Leu Leu Gln Ile Lys Thr Leu Leu Ala Xaa Ser Cys His Glu Gln Glu
```

```
20                  25                  30

Met Val Val Ser Ser Leu Val Ile Gly
    35                  40
```

(2) INFORMATION FOR SEQ ID NO: 321:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 47 amino acids
  (B) TYPE: AMINO ACID
  (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
  (A) ORGANISM: Homo Sapiens
  (F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
  (A) NAME/KEY: sig_peptide
  (B) LOCATION: -24..-1
  (C) IDENTIFICATION METHOD: Von Heijne matrix
  (D) OTHER INFORMATION: score 5.2
  seq LLFPVGRSWSCFA/QT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:

```
Met Glu Thr Phe Leu Glu Pro Asn Asn Lys Lys Leu Leu Phe Pro Val
                -20                 -15                 -10

Gly Arg Ser Trp Ser Cys Phe Ala Gln Thr Xaa Ser Leu Ala Lys Tyr
             -5                   1               5

Ile Pro Tyr Ser Leu Trp Lys Tyr Ser Val Leu Ser Gly His Ser
    10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 322:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 110 amino acids
  (B) TYPE: AMINO ACID
  (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
  (A) ORGANISM: Homo Sapiens
  (F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
  (A) NAME/KEY: sig_peptide
  (B) LOCATION: -15..-1
  (C) IDENTIFICATION METHOD: Von Heijne matrix
  (D) OTHER INFORMATION: score 5.1
  seq FLWGLALPLFFFC/WE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 322:

```
Met Gly Phe Leu Trp Gly Leu Ala Leu Pro Leu Phe Phe Phe Cys Trp
-15                 -10                 -5                   1

Glu Val Gly Val Ser Gly Ser Ser Ala Gly Pro Ser Thr Arg Arg Ala
             5                  10                  15

Asp Thr Ala Met Thr Thr Asp Asp Thr Glu Val Pro Ala Met Thr Leu
        20                  25                  30

Ala Pro Gly His Ala Ala Leu Glu Thr Gln Thr Leu Ser Ala Glu Thr
    35                  40                  45

Ser Ser Arg Ala Ser Thr Pro Ala Gly Pro Val Pro Glu Ala Glu Thr
50                  55                  60                  65

Arg Gly Ala Lys Arg Ile Ser Pro Ala Arg Glu Thr Arg Ser Phe Thr
                70                  75                  80

Lys Thr Xaa Pro Asn Phe Met Val Leu Xaa Xaa Xaa Val Thr
            85                  90                  95
```

(2) INFORMATION FOR SEQ ID NO: 323:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq WLLSDILGQGATA/NV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:

```
Met Gln Ser Thr Ser Asn His Leu Trp Leu Leu Ser Asp Ile Leu Gly
    -20                 -15                 -10

Gln Gly Ala Thr Ala Asn Val Phe Arg Gly Arg His Lys Lys Thr Gly
 -5                  1               5                  10

Asp Leu Phe Ala Ile Lys Val Phe Asn Asn Ile Ser Phe Leu Arg Pro
            15                  20                  25

Val Asp Val Gln Met Arg Glu Phe Glu Val Leu Lys Lys Leu Asn His
        30                  35                  40
```

Lys Asn Ile Val Lys Leu Phe Ala Ile Glu Glu Glu Thr Gly
45 50 55

(2) INFORMATION FOR SEQ ID NO: 324:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymphocytes

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq ICAGSVLPPYSNC/QM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 324:

Met Val Glu Ile Cys Ala Gly Ser Val Leu Pro Pro Tyr Ser Asn Cys
-15 -10 -5

Gln Met Pro Glu Pro Ser Ile Phe Thr Leu Ile His Phe His Thr Tyr
1 5 10 15

Tyr Cys Leu Thr Thr Pro Gln
20

(2) INFORMATION FOR SEQ ID NO: 325:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -43..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LLAFGTSCSVVXY/XP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:

Met Val Ala Pro Val Leu Glu Thr Ser His Val Phe Cys Cys Pro Asn
-40 -35 -30

Arg Val Arg Gly Val Leu Asn Trp Xaa Ser Gly Pro Arg Gly Leu Leu
-25 -20 -15

Ala Phe Gly Thr Ser Cys Ser Val Val Xaa Tyr Xaa Pro Leu Xaa Arg
-10 -5 1 5

Val Val Val Thr Xaa Leu Xaa Gly His Thr Ala Arg Val Asn Cys Ile
10 15 20

Gln Trp Ile Xaa Lys Gln Xaa Gly Met
25 30

(2) INFORMATION FOR SEQ ID NO: 326:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -70..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq QLLLATLQEAATT/QE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:

Met Asp Ser Leu Arg Lys Met Leu Ile Ser Val Ala Met Leu Gly Ala
-70 -65 -60 -55

Xaa Ala Gly Val Gly Tyr Ala Leu Leu Val Ile Val Thr Pro Gly Glu
-50 -45 -40

Arg Arg Lys Gln Glu Met Leu Lys Glu Met Pro Leu Gln Asp Pro Arg
-35 -30 -25

Ser Arg Glu Glu Ala Ala Arg Thr Gln Gln Leu Leu Leu Ala Thr Leu
-20 -15 -10

Gln Glu Ala Ala Thr Thr Gln Glu Asn Val Ala Trp Arg Lys Asn Trp
-5 1 5 10

Met Val Gly Gly Glu Gly Gly Ala Thr Gly Xaa His Arg Glu Thr Gly
15 20 25

Leu Ala Ser Val Gly Ala Gly Pro Trp Leu Gly Arg Arg Asn Pro Arg
30 35 40

```
Gln Leu Ser Pro Ser
        45
```

(2) INFORMATION FOR SEQ ID NO: 327:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq LLPFGMLCASSTT/KC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:

```
Met Arg Gln Thr Leu Pro Cys Ile Tyr Phe Trp Gly Gly Leu Leu Pro
    -25                 -20                 -15

Phe Gly Met Leu Cys Ala Ser Ser Thr Thr Lys Cys Thr Val Ser His
-10                 -5                   1                   5

Glu Val Ala Asp Cys Ser His Leu Lys Leu Thr Gln Val Pro Asp Asp
            10                  15                  20

Leu Pro Thr Asn Ile Thr Val Leu Asn Leu Thr His Asn Gln Leu Arg
        25                  30                  35

Arg Leu Pro Ala Ala Asn Phe Thr Arg Tyr Ser Gln Leu Thr Ser Leu
    40                  45                  50

Asp Val Gly Phe Asn Thr Ile Ser Lys Leu Glu
 55                 60                  65
```

(2) INFORMATION FOR SEQ ID NO: 328:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -110..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq HTXGLLGFGRXQG/SI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 328:

Met Ala Asp Asp Leu Glu Gln Gln Ser Gln Gly Trp Leu Ser Ser Trp
-110 -105 -100 -95

Leu Pro Thr Trp Arg Pro Thr Ser Met Ser Gln Leu Lys Asn Val Glu
-90 -85 -80

Ala Arg Ile Leu Gln Cys Leu Gln Asn Lys Phe Leu Ala Arg Tyr Val
-75 -70 -65

Ser Leu Pro Asn Gln Asn Lys Ile Trp Thr Val Thr Val Ser Pro Glu
-60 -55 -50

Gln Asn Asp Arg Thr Pro Leu Val Met Val His Gly Phe Gly Gly Gly
-45 -40 -35

Val Gly Leu Trp Ile Leu Asn Met Asp Ser Leu Xaa Ala Arg Arg Thr
-30 -25 -20 -15

Leu His Thr Xaa Gly Leu Leu Gly Phe Gly Arg Xaa Gln Gly Ser Ile
-10 -5 1

Pro Lys Gly Pro Glu Gly Leu Xaa Asp Glu Phe Val Xaa Ser Ile Xaa
5 10 15

Thr Trp Arg Glu Thr Trp
20

(2) INFORMATION FOR SEQ ID NO: 329:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Large intestine

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq PLSMILLSDKIQS/SK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

```
Met Lys Val Thr Gly Ile Thr Ile Leu Phe Trp Pro Leu Ser Met Ile
                -20                 -15                 -10

Leu Leu Ser Asp Lys Ile Gln Ser Ser Lys Arg Glu Val Gln Cys Asn
            -5                  1                   5

Phe Thr Glu Lys Asn Tyr Thr Leu Ile Pro Ala Asp Ile Lys Lys Asp
    10                  15                  20

Val Thr Ile Leu Asp Leu Ser Tyr Asn Gln Xaa Thr Leu Asn Gly Thr
25                  30                  35                  40

Asp Thr
```

(2) INFORMATION FOR SEQ ID NO: 330:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -96..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq HLSWSSSAYQAWA/QE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:

```
Met Ala Ala Gly Arg Ala Gln Val Pro Ser Ser Glu Gln Ala Trp Leu
    -95                 -90                 -85

Glu Asp Ala Gln Val Phe Ile Gln Lys Thr Leu Cys Pro Ala Val Lys
-80                 -75                 -70                 -65

Glu Pro Asn Val Gln Leu Thr Pro Leu Val Ile Asp Cys Val Lys Thr
                -60                 -55                 -50

Val Trp Leu Ser Gln Gly Arg Asn Gln Gly Ser Thr Leu Pro Leu Ser
            -45                 -40                 -35

Tyr Ser Phe Val Ser Val Gln Asp Leu Lys Thr His Gln Arg Leu Pro
        -30                 -25                 -20

Cys Cys Ser His Leu Ser Trp Ser Ser Ser Ala Tyr Gln Ala Trp Ala
    -15                 -10                 -5

Gln Glu Ala Gly Pro Asn Gly Asn Pro Pro Gly
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 331:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq STCCWCTPGGAST/ID

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:

```
Met Ser Thr Cys Cys Trp Cys Thr Pro Gly Gly Ala Ser Thr Ile Asp
                -10                 -5                   1

Phe Leu Lys Arg Tyr Ala Ser Asn Thr Pro Ser Gly Glu Phe Gln Thr
         5                  10                  15

Ala Asp Glu Asp Leu Cys Tyr Cys Leu Gly
     20                  25
```

(2) INFORMATION FOR SEQ ID NO: 332:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq VVEILPYLPCLTA/RD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:

```
Met Pro Phe Ala Glu Asp Lys Thr Tyr Lys Tyr Ile Cys Arg Asn Phe
```

```
-35                     -30                     -25

Ser Asn Phe Cys Xaa Val Asp Val Val Glu Ile Leu Pro Tyr Leu Pro
-20                 -15                 -10                 -5

Cys Leu Thr Ala Arg Asp Gln Asp Arg Leu Arg Ala Thr Cys Thr Leu
                1               5                   10

Ser Gly Asn Arg Ala
        15
```

(2) INFORMATION FOR SEQ ID NO: 333:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -107..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq IVLVLLLGRYTEE/EQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:

```
Met Ala Glu Ser Glu Asp Arg Ser Leu Arg Ile Val Leu Val Gly Lys
        -105                -100                -95

Thr Gly Ser Gly Lys Ser Ala Thr Ala Asn Thr Ile Leu Gly Glu Glu
    -90                 -85                 -80

Ile Phe Asp Ser Arg Ile Ala Ala Gln Ala Val Thr Lys Asn Cys Gln
-75                 -70                 -65                 -60

Lys Ala Ser Arg Glu Trp Gln Gly Arg Asp Leu Leu Val Val Asp Thr
                -55                 -50                 -45

Pro Gly Leu Phe Asp Thr Lys Glu Ser Leu Xaa Thr Thr Cys Lys Glu
            -40                 -35                 -30

Ile Xaa Arg Cys Ile Ile Ser Ser Cys Pro Gly Pro His Ala Ile Val
        -25                 -20                 -15

Leu Val Leu Leu Leu Gly Arg Tyr Thr Glu Glu Glu Gln Lys Thr Val
    -10                 -5                  1               5

Ala Leu Ile Xaa Leu
                10
```

(2) INFORMATION FOR SEQ ID NO: 334:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -49..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LLXCVGNFFGSTQ/DA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:

```
Met Ala Gln Lys Pro Leu Arg Leu Leu Ala Cys Gly Asp Val Glu Gly
                -45                 -40                 -35

Lys Phe Asp Ile Leu Phe Asn Arg Val Gln Ala Ile Gln Lys Xaa Ser
            -30                 -25                 -20

Gly Asn Phe Asp Leu Leu Xaa Cys Val Gly Asn Phe Phe Gly Ser Thr
        -15                 -10                 -5

Gln Asp Ala Glu Trp Glu Glu Tyr Lys Thr Gly Ile Lys Lys Ala Pro
      1               5                  10                  15

Ile Gln Thr Tyr Val Leu Gly Ala Asn Asn Gln Glu Thr Val Lys Tyr
                 20                  25                  30

Phe Gln Asp Ala Asp Gly Cys Glu Leu Ala Glu Asn Ile Thr Tyr Leu
             35                  40                  45

Gly Arg Gly
         50
```

(2) INFORMATION FOR SEQ ID NO: 335:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 109 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq RPVLLHLHQTAHA/DE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:

```
Met Glu Ser Arg Lys Asp Ile Thr Asn Gln Glu Glu Leu Trp Lys Met
        -50                 -45                 -40

Lys Pro Arg Arg Asn Leu Glu Glu Asp Asp Tyr Leu His Lys Asp Thr
    -35                 -30                 -25

Gly Glu Thr Ser Met Leu Lys Arg Pro Val Leu Leu His Leu His Gln
-20                 -15                 -10                  -5

Thr Ala His Ala Asp Glu Phe Asp Cys Pro Ser Glu Leu Gln His Thr
                 1               5                  10

Gln Gln Leu Phe Pro Gln Trp His Leu Pro Ile Lys Ile Ala Ala Ile
        15                  20                  25

Ile Ala Xaa Leu Thr Phe Leu Tyr Thr Leu Leu Arg Glu Val Xaa His
    30                  35                  40

Pro Leu Ala Thr Ser His Gln Gln Tyr Phe Tyr Lys Ile
 45                  50                  55
```

(2) INFORMATION FOR SEQ ID NO: 336:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq RPVLLHLHQTAHA/DE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

```
Met Glu Ser Arg Lys Asp Ile Thr Asn Gln Glu Glu Xaa Trp Lys Met
        -50                 -45                 -40

Lys Pro Arg Arg Asn Leu Glu Glu Asp Asp Tyr Leu His Lys Asp Thr
    -35                 -30                 -25
```

Gly Glu Thr Ser Met Leu Lys Arg Pro Val Leu Leu His Leu His Gln
-20 -15 -10 -5

Thr Ala His Ala Asp Glu Phe Asp Cys Pro Ser Glu Leu Gln His Thr
1 5 10

Gln Gly

(2) INFORMATION FOR SEQ ID NO: 337:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Colon

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq STLASVPPAATFG/AD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:

Met Ala Ala Thr Cys Glu Ile Ser Asn Ile Phe Ser Asn Tyr Phe Ser
-35 -30 -25

Ala Met Tyr Ser Ser Glu Asp Ser Thr Leu Ala Ser Val Pro Pro Ala
-20 -15 -10 -5

Ala Thr Phe Gly Ala Asp Asp Leu Val Leu Thr Leu Ser Asn Pro Gln
1 5 10

Met Ser Leu Glu Gly Thr Glu Lys Ala Ser Trp Leu Gly Glu Gln Pro
15 20 25

Gln Xaa Trp Ser Lys Thr Gln Val Leu Asp Trp Ile Ser Tyr Gln Val
30 35 40

Glu Lys Asn Lys Tyr Asp Ala Thr Gly
45 50

(2) INFORMATION FOR SEQ ID NO: 338:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 123 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -58..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq LVSFAVSSEGTEQ/GE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:

```
Met Arg Asp Cys Pro Gly Val Glu Xaa Ile Leu Asp Cys Ser Xaa Arg
            -55                 -50                 -45

Gln Lys Thr Glu Gly Cys Arg Leu Gln Ala Gly Lys Glu Cys Val Asp
        -40                 -35                 -30

Ser Pro Val Glu Gly Gly Gln Ser Glu Ala Pro Pro Ser Leu Val Ser
    -25                 -20                 -15

Phe Ala Val Ser Ser Glu Gly Thr Glu Gln Gly Glu Asp Pro Arg Ser
-10                 -5                  1               5

Glu Lys Asp His Ser Arg Pro His Lys His Arg Ala Arg His Ala Arg
            10                  15                  20

Leu Arg Arg Ser Glu Ser Leu Ser Xaa Lys Gln Val Lys Glu Ala Lys
        25                  30                  35

Ser Xaa Cys Lys Ser Ile Ala Leu Leu Leu Thr Asp Ala Pro Xaa Pro
    40                  45                  50

Asn Ser Lys Gly Val Leu Met Phe Lys Lys Arg
55                  60                  65
```

(2) INFORMATION FOR SEQ ID NO: 339:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq LVFNFLLILTILT/IW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:

```
Met Glu Arg Gln Ser Arg Val Met Ser Glu Lys Asp Glu Tyr Gln Phe
    -35                 -30                 -25

Gln His Gln Gly Ala Val Glu Leu Leu Val Phe Asn Phe Leu Leu Ile
    -20                 -15                 -10

Leu Thr Ile Leu Thr Ile Trp Leu Phe Lys Asn His Arg Phe Arg Phe
-5                      1               5                   10

Leu His Glu Thr Gly Gly Ala Met Val Tyr
            15                  20
```

(2) INFORMATION FOR SEQ ID NO: 340:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 13.4
seq SLLLVQLLTPCSA/QF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

```
Met Lys Met Ala Ser Ser Leu Ala Phe Leu Leu Leu Asn Phe His Val
                -25                 -20                 -15

Ser Leu Leu Leu Val Gln Leu Leu Thr Pro Cys Ser Ala Gln Phe Ser
            -10                 -5                  1

Val Leu Xaa Xaa Ser Gly Pro Ile Leu Ala Met Val Gly Glu Asp Ala
    5                   10                  15

Asp Leu
20
```

(2) INFORMATION FOR SEQ ID NO: 341:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 12.6
seq LLALLTVSTPSWC/QS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:

```
Met Val Phe Leu Pro Leu Lys Trp Ser Leu Ala Thr Met Ser Phe Leu
        -30                 -25                 -20

Leu Ser Ser Leu Leu Ala Leu Leu Thr Val Ser Thr Pro Ser Trp Cys
    -15                 -10                 -5

Gln Ser Thr Glu Ala Ser Pro Lys Arg
  1               5
```

(2) INFORMATION FOR SEQ ID NO: 342:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.8
seq SLLLLLLXCVHWS/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:

```
Met Glu Ser Ala Ala Ala Leu His Phe Ser Arg Pro Ala Ser Leu Leu
    -25                 -20                 -15

Leu Leu Leu Leu Xaa Cys Val His Trp Ser Gln Pro Ser Leu Leu Ser
-10                 -5                  1               5

Trp
```

(2) INFORMATION FOR SEQ ID NO: 343:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.2
seq AFLLLVALSYTLA/RD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

```
Met Glu Lys Ile Pro Val Ser Ala Phe Leu Leu Leu Val Ala Leu Ser
-20                 -15                 -10                  -5

Tyr Thr Leu Ala Arg Asp Thr Thr Val Lys Pro Gly Ala Lys Lys Asp
                 1               5                  10

Thr Lys Asp Ser Arg Pro Lys Leu Pro Gln Thr Leu Ser Arg Gly Trp
        15                  20                  25

Gly Asp Gln Leu Ile Trp Thr Gln Thr Tyr Glu Glu Ala Leu Tyr Lys
    30                  35                  40

Ser Lys Thr Ser Asn Lys Pro Leu Met Ile Ile His His Leu Asp Glu
45                  50                  55                  60

Cys Pro His Ser Gln Ala Leu Lys Lys Val Phe Ala Glu Asn Lys Glu
                65                  70                  75

Ile Gln Lys Leu Ala Glu Gln Phe Val Leu Leu Asn Leu Val Tyr Glu
            80                  85                  90

Thr Thr Asp Lys
        95
```

(2) INFORMATION FOR SEQ ID NO: 344:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -46..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.3
seq LVLLLVLTLLCSL/VP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:

Met Gly Pro Trp Gly Glu Pro Glu Leu Leu Val Trp Arg Pro Glu Ala
-45 -40 -35

Val Ala Ser Glu Pro Pro Val Pro Val Gly Leu Glu Val Lys Leu Gly
-30 -25 -20 -15

Ala Leu Val Leu Leu Leu Val Leu Thr Leu Leu Cys Ser Leu Val Pro
-10 -5 1

Ile Cys Val Leu Arg Arg Pro Gly Ala Asn His Glu Gly Ser Ala Ser
5 10 15

Arg Gln Lys Ala Leu Ser Pro Lys
20 25

(2) INFORMATION FOR SEQ ID NO: 345:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.1
seq LLLQLAVLGAALA/AA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:

Met Ala Pro Leu Leu Leu Gln Leu Ala Val Leu Gly Ala Ala Leu Ala
-15 -10 -5

Ala Ala Ala Leu Val Leu Ile Ser Ile Val Ala Phe Thr Thr Ala Thr
1 5 10 15

Lys Met Pro Ala Leu His Arg His Glu Glu Glu Lys Phe Phe Leu Asn
20 25 30

Ala Lys Gly Gln Lys Glu Thr Leu Pro Ser Ile Trp Asp Ser Pro Thr
35 40 45

Lys Gln Leu Ser Val Val Val Pro Ser Tyr Asn Glu Glu Lys Arg Leu
50 55 60

Pro Val Met Met Asp Glu Ala Leu Ser Tyr Leu Glu Lys Arg Gln Lys
65 70 75 80

Arg Asp Pro Ala Phe Thr Tyr Glu Val Ile Val Val Asp Asp Gly Ser
85 90 95

Lys Asp Gln Thr Ser Lys
100

(2) INFORMATION FOR SEQ ID NO: 346:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 127 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymphocytes

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.8
seq SALLVGFLSVIFA/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:

Met Ala Met Glu Gly Tyr Trp Arg Phe Leu Xaa Leu Leu Gly Ser Ala
-25 -20 -15

Leu Leu Val Gly Phe Leu Ser Val Ile Phe Ala Leu Val Trp Val Leu
-10 -5 1 5

His Tyr Arg Glu Gly Leu Gly Trp Asp Gly Ser Ala Leu Glu Phe Asn
10 15 20

Trp Xaa Pro Val Leu Met Val Thr Gly Phe Val Phe Ile Gln Gly Ile
25 30 35

Ala Ile Ile Val Tyr Arg Leu Pro Trp Thr Trp Lys Cys Ser Lys Leu
40 45 50

Leu Met Lys Ser Ile His Ala Xaa Leu Asn Ala Val Ala Ala Ile Leu
55 60 65

Ala Ile Ile Ser Val Val Ala Val Phe Glu Asn His Asn Val Asn Asn
70 75 80 85

Ile Ala Asn Met Tyr Ser Leu His Ser Trp Val Gly Leu Ile Ala
90 95 100

(2) INFORMATION FOR SEQ ID NO: 347:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 129 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.3
seq LALSLLILVLAFG/IP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 347:

Met Ala Gln Ser Leu Ala Leu Ser Leu Leu Ile Leu Val Leu Ala Phe
-15 -10 -5

Gly Ile Pro Arg Thr Gln Gly Ser Asp Gly Gly Ala Gln Asp Cys Cys
1 5 10 15

Leu Lys Tyr Ser Gln Arg Lys Ile Pro Ala Lys Val Val Arg Ser Tyr
20 25 30

Arg Lys Gln Glu Pro Ser Leu Gly Cys Ser Ile Pro Ala Ile Leu Phe
35 40 45

Leu Pro Arg Lys Arg Ser Gln Ala Glu Leu Cys Ala Asp Pro Lys Glu
50 55 60

Leu Trp Val Gln Gln Leu Met Gln His Leu Asp Lys Thr Pro Ser Pro
65 70 75

Gln Lys Pro Ala Gln Gly Cys Arg Lys Asp Arg Gly Ala Ser Lys Thr
80 85 90 95

Gly Lys Lys Gly Lys Gly Ser Lys Gly Cys Lys Arg Thr Glu Arg Ser
100 105 110

Gln

(2) INFORMATION FOR SEQ ID NO: 348:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 56 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.4
seq AMWLLCVALAVLA/WG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

```
Met Glu Ala Met Trp Leu Leu Cys Val Ala Leu Ala Val Leu Ala Trp
-15                 -10                 -5                   1

Gly Phe Leu Trp Val Trp Asp Ser Ser Glu Arg Met Lys Ser Arg Glu
             5                  10                  15

Gln Gly Xaa Arg Leu Gly Ala Glu Ser Arg Thr Leu Leu Val Ile Ala
        20                  25                  30

His Pro Asp Asp Glu Ala Met Trp
    35                  40
```

(2) INFORMATION FOR SEQ ID NO: 349:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -38..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.8
seq LVFTVSLFAWICC/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:

```
Met Ala Pro Ile Thr Thr Ser Arg Glu Glu Phe Asp Glu Ile Pro Thr
            -35                 -30                 -25

Val Val Gly Ile Phe Ser Ala Phe Gly Leu Val Phe Thr Val Ser Leu
        -20                 -15                 -10

Phe Ala Trp Ile Cys Cys Gln Arg Lys Ser Ser Lys Ser Asn Lys Thr
    -5                   1               5                  10
```

```
Pro Pro Tyr Lys Phe Val His Val Leu Xaa Gly Val Asp Ile Tyr Pro
                15                  20                  25

Glu Asn Leu Asn Ser Lys Lys Lys
            30
```

(2) INFORMATION FOR SEQ ID NO: 350:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 121 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq GWLVLCVLAISLA/SM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:

```
Met Glu Gly Pro Arg Gly Trp Leu Val Leu Cys Val Leu Ala Ile Ser
            -15                 -10                 -5

Leu Ala Ser Met Val Thr Glu Asp Leu Cys Arg Ala Pro Asp Gly Lys
         1               5                  10

Lys Gly Glu Ala Gly Arg Pro Gly Arg Arg Gly Arg Pro Gly Leu Lys
 15                 20                  25                  30

Gly Glu Gln Gly Glu Pro Gly Ala Pro Gly Ile Arg Thr Gly Ile Gln
                35                  40                  45

Gly Leu Lys Gly Asp Gln Gly Glu Pro Gly Pro Ser Gly Asn Pro Gly
            50                  55                  60

Lys Val Gly Tyr Pro Gly Pro Ser Gly Pro Leu Gly Ala Arg Gly Ile
        65                  70                  75

Pro Gly Ile Lys Gly Thr Lys Gly Ser Pro Gly Asn Ile Lys Asp Gln
    80                  85                  90

Pro Arg Pro Ala Phe Ser Ala Ile Arg
 95                 100
```

(2) INFORMATION FOR SEQ ID NO: 351:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -63..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.2
seq VLLTLLLIAFIFL/II

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:

```
Met Thr Ala Trp Glu Ala Met Ala Pro His Val Asn Pro Thr Leu Lys
            -60                 -55                 -50

Asp Lys Ala Leu Ser Pro Gln Gln Xaa Xaa Xaa Thr Ser Pro Ala Pro
        -45                 -40                 -35

Cys Xaa Ser Asn His His Asn Lys Lys His Leu Ile Leu Ala Phe Cys
    -30                 -25                 -20

Ala Gly Val Leu Leu Thr Leu Leu Leu Ile Ala Phe Ile Phe Leu Ile
-15                 -10                  -5                   1

Ile Lys Ser Tyr Arg Lys Tyr His Ser Lys Pro Gln Ala Pro Gly
            5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 352:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq LLCECLLLXAGYA/HD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:

```
Met Leu Cys Ser Leu Leu Leu Cys Glu Cys Leu Leu Leu Xaa Ala Gly
            -15                 -10                 -5

Tyr Ala His Asp Asp Asp Trp Ile Asp Pro Thr Asp Met Leu Asn Tyr
        1               5                   10

Asp Ala Ala Ser Gly Thr Met Arg Lys Ser Gln Ala Lys Tyr Gly Ile
15                  20                  25                  30

Ser Gly Glu Lys Asp Val Ser Pro Asp Leu Ser Cys Ala Xaa Glu Ile
                35                  40                  45

Ser Glu
```

(2) INFORMATION FOR SEQ ID NO: 353:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.9
seq LVXSLPVHCLTFA/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

```
Met Gly His Ala Met Gly Leu Val Xaa Ser Leu Pro Val His Cys Leu
                -15                 -10                 -5

Thr Phe Ala Ser Ser Ala Pro Ser Ser Pro Gln Pro Thr Arg Met Trp
            1                   5                   10

Phe Xaa Ala Gln Ala His Xaa Pro Pro Leu Ile Leu Gly Pro
    15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO: 354:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.7
seq CFSLVLLLTSIWT/TR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 354:

```
Met Ala Arg Cys Phe Ser Leu Val Leu Leu Leu Thr Ser Ile Trp Thr
    -15                 -10                 -5

Thr Arg Leu Leu Val Gln Gly Ser Leu Arg Ala Glu Glu Leu Ser Ile
 1               5                  10                  15

Gln Val Ser Cys Arg Ile Met Gly Ile Thr Leu Val Ser Lys Lys Ala
            20                  25                  30

Asn Gln Gln Leu Asn Phe Thr Glu Ala Lys Glu Ala Cys Arg Leu Leu
        35                  40                  45

Gly Leu Ser Leu Ala Gly Lys Asp Gln Val Glu Thr Ala Leu Lys Ala
    50                  55                  60

Ser Phe Glu Thr Cys Ser Tyr Gly Trp Val Gly Asp Gly Phe Val Val
65                  70                  75                  80

Ile Ser Arg Ile Ser Pro Asn Pro Lys Cys Gly Lys Asn Gly Val Gly
                85                  90                  95

Val Leu Ile Trp
            100
```

(2) INFORMATION FOR SEQ ID NO: 355:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -59..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.6
seq VLAQLAFLSQISQ/CI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:

```
Met Leu Leu Thr Arg Lys Gln Thr Cys Gln Leu Gly Ile Leu Leu Ser
                -55                 -50                 -45

Ile His Arg Gln His Ser Lys Asp Leu Gln Asp Ile Val Ala Thr Leu
            -40                 -35                 -30

Gly Pro Arg Ser Ala Thr His Pro His Gln Pro Ala Ile Gln Val Leu
        -25                 -20                 -15

Ala Gln Leu Ala Phe Leu Ser Gln Ile Ser Gln Cys Ile Ile Ser Gln
    -10                 -5                   1               5

Arg
```

(2) INFORMATION FOR SEQ ID NO: 356:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -28..-1
(C) IDENTIFICATION METHOD: Von Heijne_matrix
(D) OTHER INFORMATION: score 6.6
seq IVSLLGFVATVTL/IP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:

```
Met Trp Ala Phe Ser Glu Leu Pro Met Pro Leu Leu Ile Asn Leu Ile
            -25                 -20                 -15

Val Ser Leu Leu Gly Phe Val Ala Thr Val Thr Leu Ile Pro Ala Phe
        -10                 -5                   1

Arg Gly His Phe Ile Ala Ala Arg Leu Cys Gly Gln Asp Leu Asn Lys
  5                 10                  15                  20

Thr Ser Arg Gln Gln Ile Pro Glu Ser Gln Gly Val Ile Ser Gly Ala
                25                  30                  35

Val Phe Leu Ile Ile Leu Phe Cys
            40
```

(2) INFORMATION FOR SEQ ID NO: 357:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq PASLSLLTFKVYA/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:

```
Met Phe Lys Val Ile Gln Arg Ser Val Gly Pro Ala Ser Leu Ser Leu
            -20                 -15                 -10

Leu Thr Phe Lys Val Tyr Ala Ala Pro Lys Lys Asp Ser Pro Pro Lys
         -5                   1               5

Asn Ser Val Lys Val Asp Glu Leu Ser Leu Tyr Ser Val Pro Glu Gly
 10                  15                  20                  25

Gln Ser Lys Tyr Val Glu Glu Ala Arg Ser Gln Leu Glu Glu Ser Ile
                 30                  35                  40

Ser Gln Leu Arg His Tyr Cys Glu Pro Tyr Thr Thr Trp Cys Gln Glu
             45                  50                  55

Thr Tyr
```

(2) INFORMATION FOR SEQ ID NO: 358:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -136..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq LISVALVQGWALG/GG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:

```
Met Ala Lys Ser Leu Leu Lys Thr Ala Ser Leu Ser Gly Arg Thr Lys
    -135                -130                -125

Leu Leu His Gln Thr Gly Leu Ser Leu Tyr Ser Thr Ser His Gly Phe
-120                -115                -110                -105

Tyr Glu Glu Glu Val Lys Lys Thr Leu Gln Gln Phe Pro Gly Gly Ser
                -100                -95                 -90

Ile Asp Leu Gln Lys Glu Asp Asn Gly Ile Gly Ile Leu Thr Leu Asn
            -85                 -80                 -75

Asn Pro Ser Arg Met Asn Ala Phe Ser Gly Val Met Met Leu Gln Leu
        -70                 -65                 -60

Leu Glu Lys Val Ile Glu Leu Glu Asn Trp Thr Glu Gly Lys Gly Leu
    -55                 -50                 -45

Ile Val Arg Gly Ala Lys Asn Thr Phe Ser Ser Gly Ser Asp Leu Asn
-40                 -35                 -30                 -25

Ala Val Lys Ser Leu Gly Leu Gln Arg Leu Pro Leu Ile Ser Val Ala
                -20                 -15                 -10

Leu Val Gln Gly Trp Ala Leu Gly Gly Gly Ala Ala
            -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 359:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 64 amino acids
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: -44..-1
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 5.3
  seq PLLKILHAAGAQG/EM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:

```
Met Thr Ser Phe Ser Thr Ser Ala Gln Cys Ser Thr Ser Asp Ser Ala
                -40                 -35                 -30

Cys Arg Ile Ser Pro Gly Gln Ile Asn Xaa Val Arg Pro Lys Leu Pro
            -25                 -20                 -15

Leu Leu Lys Ile Leu His Ala Ala Gly Ala Gln Gly Glu Met Phe Thr
```

-10 -5 1

Val Lys Glu Val Met His Tyr Leu Gly Gln Tyr Ile Met Val Lys Gln
5 10 15 20

(2) INFORMATION FOR SEQ ID NO: 360:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 158 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -112..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.1
seq AFAWLGVVPLTAC/RI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:

Met Asp Thr Ala Glu Glu Asp Ile Cys Arg Val Cys Arg Ser Glu Gly
-110 -105 -100

Thr Pro Glu Lys Pro Leu Tyr His Pro Cys Val Cys Thr Gly Ser Ile
-95 -90 -85

Lys Xaa Val His Gln Glu Cys Leu Val Gln Trp Leu Lys His Ser Arg
-80 -75 -70 -65

Lys Glu Tyr Cys Glu Leu Cys Lys His Arg Phe Ala Phe Thr Pro Ile
-60 -55 -50

Tyr Ser Pro Asp Met Pro Ser Arg Leu Pro Ile Gln Asp Ile Phe Ala
-45 -40 -35

Gly Leu Val Thr Ser Ile Gly Thr Ala Ile Arg Tyr Trp Phe His Tyr
-30 -25 -20

Thr Leu Val Ala Phe Ala Trp Leu Gly Val Val Pro Leu Thr Ala Cys
-15 -10 -5

Arg Ile Tyr Lys Cys Leu Phe Thr Gly Ser Val Ser Ser Leu Leu Thr
1 5 10 15

Leu Pro Leu Asp Met Leu Ser Thr Glu Asn Leu Leu Ala Asp Cys Leu
20 25 30

Gln Gly Cys Phe Val Val Thr Cys Thr Leu Cys Ala Phe Ile
35 40 45

(2) INFORMATION FOR SEQ ID NO: 361:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq MLIMLGIFFNVHS/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

```
Met Leu Ile Met Leu Gly Ile Phe Phe Asn Val His Ser Ala Val Leu
            -10                 -5                  1

Ile Glu Asp Val Pro Phe Thr Glu Lys Asp Phe Glu Xaa Gly Pro Gln
    5                   10                  15

Asn Ile Tyr Asn Leu Tyr Glu His Gly
20                  25
```

(2) INFORMATION FOR SEQ ID NO: 362:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 114 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -112..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq AAVAVGMLXASYA/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

```
Met Gly Gly Leu Trp Arg Pro Gly Trp Arg Cys Val Pro Phe Cys Gly
        -110                -105                -100
```

```
Trp Arg Trp Ile His Pro Gly Ser Pro Thr Arg Ala Ala Glu Arg Val
    -95                 -90                 -85

Glu Pro Phe Leu Arg Pro Glu Trp Ser Gly Thr Gly Gly Ala Glu Arg
-80                 -75                 -70                 -65

Gly Leu Arg Trp Leu Gly Thr Trp Lys Arg Cys Ser Leu Arg Ala Arg
                -60                 -55                 -50

His Pro Ala Leu Gln Pro Pro Arg Arg Pro Lys Ser Ser Asn Pro Phe
            -45                 -40                 -35

Thr Arg Ala Xaa Glu Glu Glu Arg Arg Arg Xaa Asn Lys Thr Thr Leu
        -30                 -25                 -20

Thr Tyr Val Ala Ala Val Ala Val Gly Met Leu Xaa Ala Ser Tyr Ala
    -15                 -10                 -5

Ala Val
  1
```

(2) INFORMATION FOR SEQ ID NO: 363:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -39..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq SDPLCVLFLNTSG/QQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:

```
Met Ala Ala Gln Cys Val Thr Lys Val Ala Leu Asn Val Ser Cys Ala
                -35                 -30                 -25

Asn Leu Leu Asp Lys Asp Ile Gly Ser Lys Ser Asp Pro Leu Cys Val
            -20                 -15                 -10

Leu Phe Leu Asn Thr Ser Gly Gln Gln Trp Tyr Glu Val Glu Arg Thr
        -5                  1               5

Glu Arg Ile Lys Asn Cys Leu Asn Pro Gln Phe Ser Lys Thr Phe Ile
10                  15                  20                  25

Ile Asp Tyr Tyr Phe Glu Val Val Gln Lys Leu Lys Phe Gly Val Tyr
                30                  35                  40
```

Asp Ile Xaa Asn Lys Thr Ile Glu Leu Ser Asp Asp Asp Phe Leu Gly
45 50 55

(2) INFORMATION FOR SEQ ID NO: 364:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -70..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq AVLDCAFYDPTHA/WS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:

Met Thr Gly Ser Asn Glu Phe Lys Leu Asn Gln Pro Pro Glu Asp Gly
-70 -65 -60 -55

Ile Ser Ser Val Lys Phe Ser Pro Asn Thr Ser Gln Phe Leu Leu Val
-50 -45 -40

Ser Ser Trp Asp Thr Ser Val Arg Leu Tyr Asp Val Pro Ala Asn Ser
-35 -30 -25

Met Arg Leu Lys Tyr Gln His Thr Gly Ala Val Leu Asp Cys Ala Phe
-20 -15 -10

Tyr Asp Pro Thr His Ala Trp Ser Gly Gly Leu Asp His Xaa Xaa Lys
-5 1 5 10

Met His Asp Leu Asn Thr Asp Gln Glu Asn Leu Val Gly Thr His Asp
15 20 25

Ala Pro Ile Arg Cys Val Glu Tyr Cys Pro Ser
30 35

(2) INFORMATION FOR SEQ ID NO: 365:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 48 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq AHLCWCGSHCCST/CV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

```
Met Gly Lys His Leu Trp Tyr Pro Gly Gln Ala Ser Ala His Leu Cys
-25                 -20                 -15                 -10

Trp Cys Gly Ser His Cys Cys Ser Thr Cys Val Phe Glu Asp Gln Leu
                -5                   1               5

Ser Asp Glu Arg Phe Gln Arg Ser Asn Ala Pro Ser Val Asn Ser Asp
        10                  15                  20
```

(2) INFORMATION FOR SEQ ID NO: 366:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 73 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq MLAVSLTVXLLGA/MM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366:

```
Met Leu Ala Val Ser Leu Thr Val Xaa Leu Leu Gly Ala Met Met Leu
            -10                 -5                   1

Leu Glu Ser Pro Ile Asp Pro Gln Pro Leu Ser Phe Lys Glu Pro Pro
    5                   10                  15

Leu Leu Leu Gly Val Leu His Pro Asn Thr Lys Leu Arg Gln Ala Glu
20                  25                  30                  35

Arg Leu Phe Glu Asn Gln Leu Val Gly Pro Glu Ser Ile Ala His Ile
                40                  45                  50
```

```
Gly Asp Val Met Phe Thr Gly Ser Trp
            55                  60
```

(2) INFORMATION FOR SEQ ID NO: 367:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Ovary

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -76..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq MLELDLLVFHLWG/SQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

```
Met Ser Ser Thr Leu Ala Lys Ile Ala Glu Ile Glu Ala Glu Met Ala
    -75                 -70                 -65

Arg Thr Gln Lys Asn Lys Ala Thr Ala His His Leu Gly Leu Leu Lys
-60                 -55                 -50                 -45

Ala Arg Leu Ala Lys Leu Arg Arg Glu Leu Ile Thr Pro Lys Gly Gly
                -40                 -35                 -30

Gly Gly Gly Gly Pro Gly Glu Gly Phe Asp Trp Pro Arg Gln Val Met
            -25                 -20                 -15

Leu Glu Leu Asp Leu Leu Val Phe His Leu Trp Gly Ser Gln His Cys
        -10                 -5                  1

Leu Val Thr Trp Gln Gly
  5                 10
```

(2) INFORMATION FOR SEQ ID NO: 368:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal

(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -45..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 13.9
seq LVLALLLVSAALS/SV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

```
Met Ala Ala Ala Val Pro Lys Arg Met Arg Gly Pro Ala Gln Ala Lys
-45                 -40                 -35                 -30

Leu Leu Pro Gly Ser Ala Ile Gln Ala Leu Val Gly Leu Ala Arg Pro
                -25                 -20                 -15

Leu Val Leu Ala Leu Leu Leu Val Ser Ala Ala Leu Ser Ser Val Val
            -10                  -5                   1

Ser Arg Thr Asp Ser Pro Ser Pro Leu
      5                  10
```

(2) INFORMATION FOR SEQ ID NO: 369:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 13.9
seq LLSLLFLVQGAHG/RG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

```
Met Thr Pro Gln Ser Leu Leu Gln Thr Thr Leu Phe Leu Leu Ser Leu
-25                 -20                 -15                 -10

Leu Phe Leu Val Gln Gly Ala His Gly Arg Gly His Arg Glu Asp Phe
                 -5                   1                   5

Arg Phe Cys Ser Gln Arg Asn Gln Thr His Arg Ser Ser Leu His Tyr
         10                  15                  20

Lys Pro Thr Pro Xaa Leu Arg Ile Ser Ile Glu Asn Ser Glu Glu
     25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 370:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 127 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -88..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.6
seq ILLCLLLALFASG/LI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:

```
Met Met Val Val Gly Thr Gly Thr Ser Leu Ala Leu Ser Ser Leu Leu
            -85                 -80                 -75

Ser Leu Leu Leu Phe Ala Gly Met Gln Ile Tyr Ser Arg Gln Leu Ala
        -70                 -65                 -60

Ser Thr Glu Trp Leu Thr Ile Gln Gly Gly Leu Leu Gly Ser Gly Leu
    -55                 -50                 -45

Phe Val Phe Ser Leu Thr Ala Phe Asn Asn Leu Glu Asn Leu Val Phe
-40                 -35                 -30                 -25

Gly Lys Gly Phe Gln Ala Lys Ile Phe Pro Glu Ile Leu Leu Cys Leu
            -20                 -15                 -10

Leu Leu Ala Leu Phe Ala Ser Gly Leu Ile His Xaa Val Cys Val Thr
            -5                  1               5

Thr Cys Phe Ile Phe Ser Arg Val Gly Leu Tyr Tyr Ile Asn Lys Ile
    10                  15                  20

Ser Ser Thr Leu Tyr Gln Ala Ala Ala Pro Val Leu Thr Pro Ala
25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 371:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.6
seq VFCLLAVAPGAHS/QE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

```
Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Ala Pro Gly
                -15                 -10                 -5

Ala His Ser Gln Glu Gln Leu Val Gln Ser Gly Ala Glu Val Leu Lys
            1               5                   10

Pro Gly Ala Ser Val Asn Ile Ser Cys Arg Ala Ser Gly Phe Thr Phe
    15                  20                  25

Thr Asn Tyr Tyr Val His Trp Val Arg Gln Ala Pro Gly His Gly Leu
30                  35                  40                  45

Glu Trp Met Gly Val Ile Asn Pro Val Ser Gly Tyr Thr Ser Tyr Ala
                50                  55                  60

Gln Lys Leu Gln Gly Arg Leu Thr Met Thr Thr Asp Thr Ala Ala Asn
            65                  70                  75

Ile Val Tyr Met Asp Leu Ser Arg Leu Lys Ser Asp Asp Thr Ala Val
        80                  85                  90

Tyr Phe Cys Ala Lys Val Arg Cys Leu Lys Gly Ile Cys Tyr Thr Glu
    95                  100                 105

Asp Ala Leu Asp Leu Trp
110                 115
```

(2) INFORMATION FOR SEQ ID NO: 372:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -113..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.6

seq ILLCLLLALFASG/LI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 372:

Met Arg Ile Ala Asn Arg Thr Arg Phe Ser Ser Pro Phe Leu Ala Arg
-110 -105 -100

Gly Ala Gly Trp Thr His Gly Arg Gly Met Met Val Val Gly Thr Gly
-95 -90 -85

Thr Ser Leu Ala Leu Xaa Ser Leu Leu Ser Leu Leu Leu Phe Ala Gly
-80 -75 -70

Met Gln Met Tyr Ser Arg Gln Leu Ala Ser Thr Glu Trp Leu Thr Ile
-65 -60 -55 -50

Gln Gly Gly Leu Leu Gly Ser Gly Leu Phe Val Phe Ser Leu Thr Ala
-45 -40 -35

Phe Asn Asn Leu Glu Asn Leu Val Phe Gly Lys Gly Phe Gln Ala Lys
-30 -25 -20

Ile Phe Pro Glu Ile Leu Leu Cys Leu Leu Leu Ala Leu Phe Ala Ser
-15 -10 -5

Gly Leu Ile His Arg Val Cys Val Thr Thr Cys Phe Ile Phe Ser Met
1 5 10 15

Val Gly Leu Tyr Tyr Ile Asn Lys Ile Ser Ser
20 25

(2) INFORMATION FOR SEQ ID NO: 373:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Heart

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.4
seq LMSLLLVLPVVEA/VE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 373:

Met Thr Ser Val Ser Thr Gln Leu Ser Leu Val Leu Met Ser Leu Leu
-20 -15 -10

Leu Val Leu Pro Val Val Glu Ala Val Glu Ala Gly Asp Ala Ile Ala

-5 1 5

Leu Leu Leu Gly Val Val Leu Ser Ile Thr Gly Ile Val Pro Ala Trp
10 15 20

Gly Tyr Met His Gly
25

(2) INFORMATION FOR SEQ ID NO: 374:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11.2
seq ILVVLMGLPLAQA/LD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 374:

Met Thr Pro Leu Leu Thr Leu Ile Leu Val Val Leu Met Gly Leu Pro
-20 -15 -10 -5

Leu Ala Gln Ala Leu Asp Cys His Val Cys Xaa Tyr Asn Gly Asp Asn
1 5 10

Cys

(2) INFORMATION FOR SEQ ID NO: 375:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 11

seq LLALSLLVLWTSP/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 375:

```
Met Ala Leu Leu Leu Ala Leu Ser Leu Leu Val Leu Trp Thr Ser Pro
    -15                 -10                 -5

Ala Pro Thr Leu Ser Gly Thr Asn Asp Ala Glu Asp Cys Cys Leu Ser
 1               5                  10                  15

Val Thr Gln Lys Pro Ile Pro Gly Tyr Ile Val Arg Asn Phe His Tyr
            20                  25                  30

Leu Leu Ile Lys Asp Gly Cys Arg Val Pro Ala Val Val Phe Thr Thr
        35                  40                  45

Leu Arg Gly Arg Gln Leu Cys Ala Pro Pro Asp Gln Pro Trp Val Glu
    50                  55                  60

Arg Ile Ile Gln Arg Leu Gln Arg Thr Ser Ala Lys Met Lys Xaa Arg
65                  70                  75                  80

Ser Ser Xaa Pro Met Xaa Val Xaa Arg Glu Pro Glu Ser Glu Ser Ser
                85                  90                  95

Ile Val Asn Xaa Tyr Leu Xaa Gly Glu Arg Xaa Arg
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 376:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10.5
seq RLLLLPLLLAVSG/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 376:

```
Met Gly Gly Leu Glu Pro Cys Ser Arg Leu Leu Leu Leu Pro Leu Leu
    -20                 -15                 -10

Leu Ala Val Ser Gly Leu Arg Pro Val Gln Ala Gln Ala Gln Ser Asp
-5                   1                   5                  10

Cys Ser Cys Ser Thr Val Ser Pro Gly Val Leu Ala Gly Ile Val Met
```

15 20 25

Gly Asp Leu Val Leu Thr Val Leu Ile Ala Leu Ala Val Tyr Phe Leu
30 35 40

Gly Arg Leu Val Pro Arg Gly Arg Gly Ala Ala Glu Ala Xaa Thr Arg
45 50 55

Lys Gln Arg Ile Thr Glu Thr Gly Ser Pro Tyr Gln Glu Leu Gln Gly
60 65 70 75

Gln Arg Ser Asp Val Tyr Ser
80

(2) INFORMATION FOR SEQ ID NO: 377:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 10
seq LCRALCLFPRVFA/AE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 377:

Met Glu Val Pro Pro Pro Ala Pro Arg Ser Phe Leu Cys Arg Ala Leu
-20 -15 -10

Cys Leu Phe Pro Arg Val Phe Ala Ala Glu Ala Val Thr Ala Asp Ser
-5 1 5

Glu Val Leu Glu Glu Arg Gln Lys Arg Leu Pro Tyr Xaa Pro Glu Pro
10 15 20

Tyr Tyr Arg Asn Leu Asp Gly Thr Ala Ser Gly Ser Cys Xaa Ala Lys
25 30 35 40

Met Asn Ser Arg Glu Phe Gln Arg Thr Leu Leu Ile Ser Val Arg Arg
45 50 55

Gln Leu

(2) INFORMATION FOR SEQ ID NO: 378:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.5
seq LMCLSLCTAFALS/KP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 378:

```
Met Asp Leu Arg Gln Phe Leu Met Cys Leu Ser Leu Cys Thr Ala Phe
                -15                 -10                  -5

Ala Leu Ser Lys Pro Thr Glu Lys Lys Asp Arg Val His His Glu Pro
             1               5                  10

Gln Leu Ser Asp Lys Val His Asn Asp Ala Gln Ser Phe Xaa Tyr Asp
     15                  20                  25

His Asp Ala Phe Leu Gly Ala Glu Glu Ala Lys Xaa Phe Asp Gln Leu
 30                  35                  40                  45

Thr Pro Glu Glu Ser Lys Glu Arg Leu Gly Lys Ile Val Ser Lys Ile
                 50                  55                  60

Asp Gly Asp Lys Asp Gly Phe Val Thr Val Asp Glu Leu Lys
             65                  70                  75
```

(2) INFORMATION FOR SEQ ID NO: 379:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -30..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.5
seq LLFLSQFCILSGG/ES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 379:

```
Met Ala Gly Gly Val Arg Pro Leu Arg Gly Leu Arg Ala Leu Cys Arg
-30                 -25                 -20                 -15

Val Leu Leu Phe Leu Ser Gln Phe Cys Ile Leu Ser Gly Gly Glu Ser
                -10                 -5                   1

Thr Glu Ile Pro Pro Tyr Val Met Lys Cys Pro Ser Asn Gly Leu Cys
         5                  10                  15

Ser Arg Leu Pro Ala Asp Cys Ile Asp Ser Thr Thr Asn Phe Ser Cys
     20                 25                  30

Thr Tyr Gly Lys Pro Val Thr Phe Asp Cys Xaa Val Lys Pro Ser Val
 35                 40                  45                  50

Thr Cys Val Asp Gln Asp Phe Lys Ser Gln Lys Xaa Phe Ile Ile Asn
                55                  60                  65

Met Thr Cys
```

(2) INFORMATION FOR SEQ ID NO: 380:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.3
seq VLPVILLLLGAHP/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 380:

```
Met Ala Ala Ala Ala Trp Leu Gln Val Leu Pro Val Ile Leu Leu Leu
    -20                 -15                 -10

Leu Gly Ala His Pro Ser Pro Leu Ser Phe Phe Ser Ala Gly Pro Ala
 -5                  1                   5                  10

Thr Val Ala Ala Ala Asp Arg Ser Lys Trp His Ile Pro Ile Pro Ser
             15                  20                  25

Gly Lys Asn Tyr Phe Ser Phe Gly Xaa Ile Leu Phe Arg Asn Thr Thr
        30                  35                  40

Ile Phe Leu Lys Phe Asp Gly Glu Pro Cys Asp Leu Ser Leu Asn Ile
```

45 50 55

Xaa Trp Tyr Leu Lys Ser Ala Asp Cys Tyr Asn Glu Ile Tyr Asn Phe
60 65 70 75

Lys Ala Glu Glu Val Glu Leu Tyr Leu Glu Lys Leu Lys Glu Lys Arg
80 85 90

Gly Leu Ser Gly Lys Trp
95

(2) INFORMATION FOR SEQ ID NO: 381:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.3
seq LLWLALACSPVHT/XL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 381:

Met Arg Thr Leu Phe Asn Leu Leu Trp Leu Ala Leu Ala Cys Ser Pro
-15 -10 -5

Val His Thr Xaa Leu Ser Lys Ser Asp Ala Xaa Lys Pro Pro Arg
1 5 10

(2) INFORMATION FOR SEQ ID NO: 382:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 9.3
seq LFVAIFAVPLILG/QE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 382:

```
Met Asp Val Leu Phe Val Ala Ile Phe Ala Val Pro Leu Ile Leu Gly
    -15                 -10                 -5

Gln Glu Tyr Glu Asp Glu Glu Arg Leu Gly Glu Asp Glu Tyr Tyr Gln
 1              5                   10                  15

Val Val Tyr Tyr Tyr Thr Val Thr Pro Ile Met Met Xaa Leu Gly Xaa
            20                  25                  30

Xaa Phe Thr Ile Asp Tyr Xaa Ile Phe Glu Ser Glu
        35                  40
```

(2) INFORMATION FOR SEQ ID NO: 383:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 108 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.3
seq VLPVILLLLGAHP/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 383:

```
Met Ala Ala Ala Ala Trp Leu Gln Val Leu Pro Val Ile Leu Leu Leu
    -20                 -15                 -10

Leu Gly Ala His Pro Ser Pro Leu Ser Phe Phe Ser Ala Gly Pro Ala
 -5                  1              5                   10

Thr Val Ala Ala Ala Asp Arg Ser Lys Trp His Ile Pro Ile Pro Ser
            15                  20                  25

Gly Lys Asn Tyr Phe Ser Phe Gly Lys Ile Leu Phe Arg Asn Thr Thr
        30                  35                  40

Ile Phe Leu Lys Phe Asp Gly Glu Pro Cys Asp Leu Ser Leu Asn Ile
    45                  50                  55

Thr Trp Tyr Leu Lys Ser Ala Asp Cys Tyr Asn Glu Ile Tyr Asn Phe
 60                 65                  70                  75
```

Lys Ala Glu Glu Val Glu Leu Tyr Leu Glu Lys Leu
80 85

(2) INFORMATION FOR SEQ ID NO: 384:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9.2
seq LLXLALACSPVHT/TL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 384:

Met Arg Thr Leu Phe Asn Leu Leu Xaa Leu Ala Leu Ala Cys Ser Pro
-15 -10 -5

Val His Thr Thr Leu Ser Lys Ser Asp Ala Lys Lys Ala Ala Ser Lys
1 5 10

Thr Leu Leu Glu Lys Ser Gln Phe Ser Asp Lys Pro Val Gln Asp Arg
15 20 25

Gly Leu Val Val Thr Asp Gly
30 35

(2) INFORMATION FOR SEQ ID NO: 385:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -40..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 9

seq LLCLLHFSIVSVA/AX

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 385:

```
Met Gly Ser Lys Val Ala Asp Leu Leu Tyr Trp Lys Asp Thr Arg Thr
-40                 -35                 -30                 -25

Ser Gly Val Val Phe Thr Gly Leu Met Val Ser Leu Leu Cys Leu Leu
                -20                 -15                 -10

His Phe Ser Ile Val Ser Val Ala Ala Xaa Phe Gly Xaa Xaa Xaa Xaa
            -5                  1                   5

Xaa Gly Xaa Gln Ser Ser Xaa Arg Val Tyr Ala Lys Cys Cys Arg Pro
    10                  15                  20

Cys Thr Gly Gly Met Glu
25                  30
```

(2) INFORMATION FOR SEQ ID NO: 386:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.9
seq ALLIVCDVPSASA/QR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 386:

```
Met Ala Ala Arg Trp Arg Phe Trp Cys Val Ser Val Thr Met Val Val
                -25                 -20                 -15

Ala Leu Leu Ile Val Cys Asp Val Pro Ser Ala Ser Ala Gln Arg Lys
            -10                 -5                  1

Lys Glu Met Val Leu Ser Glu Lys Val Ser Gln Leu Met Glu Trp Thr
    5                   10                  15

Asn Lys Arg Pro Val Ile Arg Met Asn Gly Asp Lys Phe Arg Arg Leu
20                  25                  30                  35

Val Lys Pro His Met
                40
```

(2) INFORMATION FOR SEQ ID NO: 387:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 137 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.8
seq SAVLSGFVLGALA/FQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 387:

```
Met Glu Gly Glu Ser Thr Ser Ala Val Leu Ser Gly Phe Val Leu Gly
                -15                 -10                 -5

Ala Leu Ala Phe Gln His Leu Asn Thr Asp Ser Asp Thr Glu Gly Phe
            1               5                   10

Leu Leu Gly Glu Val Lys Gly Glu Ala Lys Asn Ser Ile Thr Asp Ser
    15                  20                  25

Gln Met Asp Asp Val Glu Val Val Tyr Thr Ile Asp Ile Gln Lys Tyr
 30                 35                  40                  45

Ile Pro Cys Tyr Gln Leu Phe Ser Phe Tyr Asn Ser Ser Gly Glu Val
                50                  55                  60

Asn Glu Gln Ala Leu Lys Lys Ile Leu Ser Asn Val Lys Lys Asn Val
            65                  70                  75

Val Gly Trp Tyr Lys Phe Arg Arg His Ser Asp Gln Ile Met Thr Phe
        80                  85                  90

Arg Glu Arg Leu Leu His Lys Asn Leu Gln Glu His Phe Ser Asn Gln
    95                  100                 105

Asp Leu Val Phe Leu Leu Leu Thr Pro
110                 115
```

(2) INFORMATION FOR SEQ ID NO: 388:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.8
seq VPMLLLIVGGSFG/LR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 388:

```
Met Phe Ala Pro Ala Val Met Arg Ala Phe Arg Lys Asn Lys Thr Leu
        -30                 -25                 -20

Gly Tyr Gly Val Pro Met Leu Leu Leu Ile Val Gly Gly Ser Phe Gly
    -15                 -10                 -5

Leu Arg Glu Phe Ser Gln Ile Arg Tyr Asp Ala Val Lys Ser Lys Met
  1               5                  10                  15

Asp Pro Glu Arg
             20
```

(2) INFORMATION FOR SEQ ID NO: 389:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -136..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.7
seq AVALSLFLGWLGA/DR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 389:

```
Met Ala Ala Ala Trp Xaa Ser Gly Pro Ser Ala Pro Glu Ala Val Thr
    -135                -130                -125

Ala Arg Leu Val Gly Val Leu Trp Phe Val Ser Val Thr Thr Gly Pro
-120                -115                -110                -105

Trp Gly Ala Val Ala Thr Ser Ala Gly Gly Glu Glu Ser Leu Lys Cys
                -100                -95                 -90
```

```
Glu Asp Leu Lys Val Gly Gln Tyr Ile Cys Lys Asp Pro Lys Ile Asn
            -85                 -80                 -75

Asp Ala Thr Gln Glu Pro Val Asn Cys Thr Asn Tyr Thr Ala His Val
        -70                 -65                 -60

Ser Cys Phe Pro Ala Pro Asn Ile Thr Cys Lys Asp Ser Ser Gly Asn
    -55                 -50                 -45

Glu Thr His Phe Thr Gly Asn Glu Val Gly Phe Phe Lys Pro Ile Ser
-40                 -35                 -30                 -25

Cys Arg Asn Val Asn Gly Tyr Ser Tyr Lys Val Ala Val Ala Leu Ser
                -20                 -15                 -10

Leu Phe Leu Gly Trp Leu Gly Ala Asp Arg Phe
            -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 390:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.6
seq LLWLALACSPVHT/TL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 390:

```
Met Arg Thr Leu Phe Asn Leu Leu Trp Leu Ala Leu Ala Cys Ser Pro
                -15                 -10                 -5

Val His Thr Thr Leu Ser Lys Ser Asp Ala Lys Lys Ala Thr Ser Gly
            1                   5                   10
```

(2) INFORMATION FOR SEQ ID NO: 391:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -42..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.6
seq ASLFLLLSLTVFS/IV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 391:

```
Met Asp Gly Gln Lys Lys Asn Trp Lys Asp Lys Val Val Asp Leu Leu
        -40                 -35                 -30

Tyr Trp Arg Asp Ile Lys Lys Thr Gly Val Val Phe Gly Ala Ser Leu
    -25                 -20                 -15

Phe Leu Leu Leu Ser Leu Thr Val Phe Ser Ile Val Ser Val Thr Ala
-10                 -5                  1                   5

Tyr Ile Ala Leu Ala Leu Leu Ser Val Thr Ile Ser Phe Arg Ile Tyr
            10                  15                  20

Lys Gly Val Ile Gln Ala Ile Gln Lys Ser Asp Glu Gly His Pro Phe
        25                  30                  35

Arg Ala Tyr Leu Glu Ser Glu Val Ala Ile Ser
    40                  45
```

(2) INFORMATION FOR SEQ ID NO: 392:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 145 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.5
seq LVLGLVLPLILWA/DR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 392:

```
Met Val Ala Pro Gly Leu Val Leu Gly Leu Val Leu Pro Leu Ile Leu
            -15                 -10                 -5

Trp Ala Asp Arg Ser Ala Gly Ile Gly Phe Arg Phe Ala Ser Tyr Ile
        1                   5                   10
```

Asn Asn Asp Met Val Leu Gln Lys Glu Pro Ala Gly Ala Val Ile Trp
15 20 25 30

Gly Phe Gly Thr Pro Gly Ala Thr Val Thr Val Thr Leu Arg Gln Gly
35 40 45

Gln Glu Thr Ile Met Lys Lys Val Thr Ser Val Lys Ala His Ser Asp
50 55 60

Thr Trp Met Val Val Leu Asp Pro Met Lys Pro Gly Gly Xaa Phe Glu
65 70 75

Val Met Ala Gln Gln Thr Leu Glu Lys Ile Asn Phe Thr Leu Arg Val
80 85 90

His Asp Val Leu Phe Gly Asp Val Trp Leu Cys Ser Gly Gln Ser Asn
95 100 105 110

Met Gln Met Thr Ala Arg Val Phe Arg Trp Arg His Val Xaa Gly Leu
115 120 125

Leu

(2) INFORMATION FOR SEQ ID NO: 393:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Ovary

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.5
seq LLTIVGLILPTRG/QT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 393:

Met Ser Pro Ser Gly Arg Leu Cys Leu Leu Thr Ile Val Gly Leu Ile
-20 -15 -10

Leu Pro Thr Arg Gly Gln Thr Leu Lys Asp Thr Thr Ser Ser Ser Ser
-5 1 5 10

Ala Asp Ser Thr Ile Met Asp Ile Gln Val Pro Thr Arg Ala Pro Asp
15 20 25

Ala Val Tyr Thr Glu Leu Gln Pro Thr Ser Pro Thr Pro Thr Trp Pro
30 35 40

```
Ala Asp Glu Thr Pro Gln Pro Gln Thr Gln Thr Gln Gln Leu Glu Gly
     45                  50                  55

Thr Asp Gly Pro Leu Val Thr Asp Pro Glu Thr Pro Arg
 60                  65                  70
```

(2) INFORMATION FOR SEQ ID NO: 394:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 114 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -47..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.3
seq LALSSLLSLLLFA/GM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 394:

```
Met Arg Ile Ala Asn Arg Thr Arg Phe Ser Leu Pro Phe Leu Ala Arg
        -45                 -40                 -35

Gly Ala Gly Trp Thr His Gly Arg Gly Met Met Val Val Gly Thr Gly
    -30                 -25                 -20

Thr Ser Leu Ala Leu Ser Ser Leu Leu Ser Leu Leu Leu Phe Ala Gly
-15                 -10                  -5                   1

Met Gln Met Tyr Ser Arg Gln Leu Ala Ser Thr Glu Trp Leu Thr Ile
             5                  10                  15

Gln Gly Gly Leu Leu Gly Ser Gly Leu Phe Val Phe Ser Leu Thr Ala
         20                  25                  30

Phe Asn Asn Leu Glu Asn Leu Val Phe Gly Lys Gly Phe Gln Ala Lys
     35                  40                  45

Ile Phe Pro Glu Ile Leu Leu Cys Leu Leu Leu Ala Leu Phe Ala Ser
 50                  55                  60                  65

Gly Pro
```

(2) INFORMATION FOR SEQ ID NO: 395:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 amino acids
(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.3
seq NLLLLHCVSRSHS/QN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 395:

```
Met Val Leu Gly Gly Cys Pro Val Ser Tyr Leu Leu Leu Cys Gly Gln
-35                 -30                 -25                 -20

Ala Ala Leu Leu Leu Gly Asn Leu Leu Leu Leu His Cys Val Ser Arg
                -15                 -10                  -5

Ser His Ser Gln Asn Ala Thr Ala Glu Pro Glu Leu Thr Ser Ala Gly
              1               5                  10

Ala Ala Gln Pro Glu Gly Pro Gly Gly Ala Ala Ser Trp Glu Tyr Gly
     15                  20                  25

Asp Pro His Ser Pro Val Ile Leu Xaa Ser Tyr Leu Pro Asp Glu Phe
 30                  35                  40                  45

Ile Glu Cys Glu Asp Arg
                 50
```

(2) INFORMATION FOR SEQ ID NO: 396:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -53..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8.1
seq IYALFLLVGVCVA/CV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 396:

```
Met Gly Ser Val Leu Gly Leu Cys Ser Met Ala Ser Trp Ile Pro Cys
            -50                 -45                 -40

Leu Cys Gly Ser Ala Pro Cys Leu Leu Cys Arg Cys Cys Pro Ser Gly
        -35                 -30                 -25

Asn Asn Ser Thr Val Thr Arg Leu Ile Tyr Ala Leu Phe Leu Leu Val
    -20                 -15                 -10

Gly Val Cys Val Ala Cys Val Met Leu Ile Pro Gly Met Glu Glu Gln
 -5                  1               5                  10

Leu Asn Lys Ile Pro Gly Phe Cys Glu Asn Glu Lys Gly Val Val Pro
            15                  20                  25

Cys Asn Ile Leu Val Gly Tyr Lys Ala Val Tyr Arg Leu Cys Phe Gly
        30                  35                  40

Leu Ala Met Xaa Tyr Leu Leu Leu Ser Leu Leu Met Ile Lys Val Lys
    45                  50                  55

Ser Ser Ser Asp Pro Arg Ala Ala Val His Asn Gly Phe
60                  65                  70
```

(2) INFORMATION FOR SEQ ID NO: 397:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 155 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -57..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 8
seq IVRLVAFCPFASS/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 397:

```
Met Val Leu Leu His Val Leu Phe Glu His Ala Val Gly Tyr Ala Leu
    -55                 -50                 -45

Leu Ala Leu Lys Glu Val Glu Glu Ile Ser Leu Leu Gln Pro Gln Val
    -40                 -35                 -30

Glu Glu Ser Val Leu Asn Leu Gly Lys Phe His Ser Ile Val Arg Leu
-25                 -20                 -15                 -10

Val Ala Phe Cys Pro Phe Ala Ser Ser Gln Val Ala Leu Glu Asn Ala
                -5                   1                   5
```

```
Asn Ala Val Ser Glu Gly Val Val His Glu Asp Leu Arg Leu Leu Leu
        10                  15                  20

Glu Thr His Leu Pro Ser Lys Lys Lys Lys Val Leu Leu Gly Val Gly
    25                  30                  35

Asp Pro Lys Ile Gly Ala Ala Ile Gln Glu Glu Leu Gly Tyr Asn Cys
40                  45                  50                  55

Gln Thr Gly Gly Val Ile Ala Glu Ile Leu Arg Xaa Val Arg Leu His
                60                  65                  70

Phe His Asn Leu Val Lys Gly Ser Asp Arg Cys Xaa Gln Leu Val Lys
            75                  80                  85

His Ser Trp Gly Trp Asp Thr Ala Ile Pro Met
        90                  95
```

(2) INFORMATION FOR SEQ ID NO: 398:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 62 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Colon

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -47..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.9
seq LLLPRVLLTMASG/SP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 398:

```
Met Ser Gly Gly Arg Ala Pro Ala Val Leu Leu Gly Gly Val Ala Ser
        -45                 -40                 -35

Leu Leu Leu Ser Phe Val Trp Met Pro Ala Leu Leu Pro Val Ala Ser
    -30                 -25                 -20

Arg Leu Leu Leu Leu Pro Arg Val Leu Leu Thr Met Ala Ser Gly Ser
-15                 -10                 -5                  1

Pro Pro Thr Gln Pro Ser Pro Ala Ser Asp Ser Gly Ile Gly
            5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 399:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.8
seq LVGFILFLTRSRG/RA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 399:

Met Val Ala Pro Val Trp Tyr Leu Val Ala Ala Ala Leu Leu Val Gly
-25 -20 -15

Phe Ile Leu Phe Leu Thr Arg Ser Arg Gly Arg Ala Ala Ser Ala Gly
-10 -5 1 5

Gln Glu Pro Leu His Asn Glu Glu Pro Gly
10 15

(2) INFORMATION FOR SEQ ID NO: 400:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.6
seq FLLVRKLPPLCHG/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 400:

Met Ala Val Leu Ala Pro Leu Ile Ala Leu Val Tyr Ser Val Pro Arg
-45 -40 -35

Leu Ser Arg Trp Leu Ala Gln Pro Tyr Tyr Leu Leu Ser Ala Leu Leu
-30 -25 -20

Ser Ala Ala Phe Leu Leu Val Arg Lys Leu Pro Pro Leu Cys His Gly

```
    -15                 -10                  -5

Leu Pro Thr Gln Xaa Glu Asp Gly Asn Pro Cys Asp Phe Asp Trp Arg
 1               5                  10                  15

Glu Val Glu Ile Leu Met Phe Leu Ser Ala Ile Val Met Met Lys Asn
            20                  25                  30

Arg Arg Ser Ile Thr Val Glu Gln His Ile Gly Asn Ile Phe Met Phe
        35                  40                  45

Ser Lys Val Ala Asn Thr Ile Leu Phe Phe Arg Leu Asp Ile Arg Met
    50                  55                  60

Gly Leu Leu Tyr Ile Thr Leu Cys Ile Val Phe Leu Met Thr Cys Lys
65                  70                  75                  80

Pro Pro Leu
```

(2) INFORMATION FOR SEQ ID NO: 401:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 148 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -69..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.6
seq FLLVRKLPPLCHG/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 401:

```
Met Val Gly Glu Ala Gly Arg Asp Leu Arg Arg Arg Arg Xaa Xaa Ala
                -65                 -60                 -55

Val Thr Ala Xaa Lys Met Ala Val Leu Ala Pro Leu Ile Ala Leu Val
            -50                 -45                 -40

Tyr Ser Val Pro Arg Leu Ser Arg Trp Leu Ala Gln Pro Tyr Tyr Leu
        -35                 -30                 -25

Leu Ser Ala Leu Leu Ser Ala Ala Phe Leu Leu Val Arg Lys Leu Pro
    -20                 -15                 -10

Pro Leu Cys His Gly Leu Pro Thr Gln Arg Glu Asp Gly Asn Pro Cys
-5                   1               5                  10

Asp Phe Asp Trp Arg Glu Val Glu Ile Leu Met Phe Leu Ser Ala Ile
            15                  20                  25
```

```
Val Met Met Lys Asn Arg Arg Ser Ile Thr Val Glu Gln His Ile Ala
        30                  35                      40

Asn Ile Phe Met Phe Ser Lys Val Ala Asn Thr Ile Leu Phe Phe Arg
    45                  50                  55

Leu Asp Ile Arg Met Gly Leu Leu Tyr Ile Thr Leu Cys Ile Val Phe
60                  65                  70                  75

Leu Met Thr Cys
```

(2) INFORMATION FOR SEQ ID NO: 402:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 139 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.6
seq LLMLLLFLSELQY/YL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 402:

```
Met Glu Ala Leu Gly Lys Leu Lys Gln Phe Asp Ala Tyr Pro Lys Thr
            -45                 -40                 -35

Leu Glu Asp Phe Arg Val Lys Thr Cys Gly Gly Ala Thr Val Thr Ile
        -30                 -25                 -20

Val Ser Gly Leu Leu Met Leu Leu Leu Phe Leu Ser Glu Leu Gln Tyr
    -15                 -10                 -5

Tyr Leu Thr Thr Glu Val His Pro Glu Leu Tyr Val Asp Lys Ser Arg
1               5                   10                  15

Gly Asp Lys Leu Lys Ile Asn Ile Asp Val Leu Phe Pro His Met Pro
            20                  25                  30

Cys Ala Tyr Leu Ser Ile Asp Ala Met Asp Val Ala Gly Glu Gln Gln
        35                  40                  45

Leu Asp Val Glu His Asn Leu Phe Lys Gln Arg Leu Asp Lys Asp Gly
    50                  55                  60

Ile Pro Val Ser Ser Glu Ala Glu Arg His Glu Leu Gly Lys Val Glu
65                  70                  75                  80
```

Val Thr Val Phe Asp Pro Asp Ser Leu Asp Pro
85 90

(2) INFORMATION FOR SEQ ID NO: 403:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 144 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.6
seq FLLVRKLPPLCHG/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 403:

Met Ala Val Leu Ala Pro Leu Ile Ala Leu Val Tyr Ser Val Pro Arg
-45 -40 -35

Leu Ser Arg Trp Leu Ala Gln Pro Tyr Tyr Leu Leu Ser Ala Leu Leu
-30 -25 -20

Ser Ala Ala Phe Leu Leu Val Arg Lys Leu Pro Pro Leu Cys His Gly
-15 -10 -5

Leu Pro Thr Gln Arg Glu Asp Gly Asn Xaa Cys Asp Phe Asp Trp Arg
1 5 10 15

Glu Val Glu Ile Leu Met Phe Leu Ser Ala Ile Val Met Met Lys Asn
20 25 30

Arg Arg Ser Ile Thr Val Glu Gln His Ile Gly Asn Ile Phe Met Phe
35 40 45

Ser Lys Val Ala Asn Thr Ile Leu Phe Phe Arg Leu Asp Ile Arg Met
50 55 60

Gly Leu Leu Xaa Ile Thr Leu Cys Ile Val Phe Leu Met Thr Cys Lys
65 70 75 80

Pro Pro Leu Tyr Met Gly Pro Glu Tyr Ile Xaa Tyr Phe Asn Asp Lys
85 90 95

(2) INFORMATION FOR SEQ ID NO: 404:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids

(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq PMLLRALAQAARA/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 404:

```
Met Arg Cys Leu Thr Thr Pro Met Leu Leu Arg Ala Leu Ala Gln Ala
                -15                 -10                 -5

Ala Arg Ala Gly Pro Pro Gly Gly Arg Ser Leu His Ser Ser Ala Val
             1               5                  10

Ala Ala Thr Tyr Lys Tyr Val Asn Met Gln Asp Pro Glu Met Asp Met
    15                  20                  25

Lys Ser Val Thr Asp Arg Ala Ala Arg Thr Leu Leu Trp Thr Glu Leu
30                  35                  40                  45

Phe Arg Gly Leu Gly Met Thr Leu Ser Tyr Leu Phe Arg Glu Pro Ala
                50                  55                  60

Thr Ile Asn Tyr Pro Phe Glu Lys Gly Pro Leu Ser Pro Arg Phe Arg
            65                  70                  75

Gly Glu His Ala Leu Arg Arg Tyr Pro Ser Gly Glu Glu Arg Cys Ile
        80                  85                  90

Ala Cys Lys Leu Cys
    95
```

(2) INFORMATION FOR SEQ ID NO: 405:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq PMLLRALAQAARA/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 405:

```
Met Arg Cys Leu Thr Thr Pro Met Leu Leu Arg Ala Leu Ala Gln Ala
                -15                 -10                 -5

Ala Arg Ala Gly Pro Pro Gly Gly Arg Ser Leu His Ser Ser Ala Val
             1               5                  10

Ala Ala Thr Tyr Lys Tyr Val Asn Met Gln Asp Pro Glu Met Asp Met
     15                  20                  25

Lys Ser Val Thr Asp Arg Ala Ala Arg Thr Leu Leu Trp Thr Glu Leu
 30                  35                  40                  45

Phe Arg Gly Leu Gly Met Thr Leu Ser Tyr Leu Phe Arg Glu Pro Xaa
                 50                  55                  60

Thr Ile Asn Tyr Pro Phe Glu Lys Gly Pro Leu Ser Pro Arg Phe Arg
             65                  70                  75

Gly Glu His Ala Leu Arg Arg Tyr Pro Ser Gly Glu Glu Arg Cys Ile
         80                  85                  90

Ala Cys Lys Leu Cys Glu Ala Ile Cys Pro Ala Gln Ala Ile Thr Ile
     95                 100                 105

Glu Ala Glu
110
```

(2) INFORMATION FOR SEQ ID NO: 406:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq ILPLLFGCLGVFG/LF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 406:

```
Met Asp Phe Ile Thr Ser Thr Ala Ile Leu Pro Leu Leu Phe Gly Cys
    -20                 -15                 -10
```

```
Leu Gly Val Phe Gly Leu Phe Arg Leu Leu Gln Trp Val Arg Gly Lys
 -5                  1               5                  10

Ala Tyr Leu Arg Asn Ala Val Val Val Ile Thr Gly Ala Thr Ser Gly
            15                  20                  25

Leu Gly Lys Glu Cys Ala Lys Val Phe Tyr Ala Ala Gly Ala Lys Leu
        30                  35                  40

Val Leu Cys Gly Arg Asn Gly Gly Ala Leu Glu Glu Leu Ile Arg Glu
    45                  50                  55

Leu Thr Ala Ser His Ala Thr Lys Val Gln Thr His Lys Pro Tyr Leu
 60                  65                  70                  75

Val Xaa Xaa Asp Leu Thr Asp Ser Gly Ala Ile Val Ala Ala Ala Ala
                80                  85                  90

Glu Ile Cys Ser
            95
```

(2) INFORMATION FOR SEQ ID NO: 407:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 47 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq LLLVTWVFTPVTT/EI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 407:

```
Met His Pro Ala Val Phe Leu Ser Leu Pro Asp Leu Arg Cys Ser Leu
                -25                 -20                 -15

Leu Leu Leu Val Thr Trp Val Phe Thr Pro Val Thr Thr Glu Ile Thr
            -10                 -5                   1

Ser Leu Asp Thr Glu Xaa Ile Asp Glu Ile Leu Asn Asn Ala Leu
     5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 408:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 63 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq LVFCVGLLTMAKA/ES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 408:

```
Met Ala Ser Leu Gly His Ile Leu Val Phe Cys Val Gly Leu Leu Thr
-20                 -15                 -10                  -5

Met Ala Lys Ala Glu Ser Pro Lys Glu His Asp Pro Phe Thr Tyr Asp
                 1                   5                  10

Tyr Gln Ser Leu Gln Ile Gly Gly Leu Val Ile Ala Gly Ile Leu Phe
         15                  20                  25

Ile Leu Gly Ile Leu Ile Val Leu Ser Arg Arg Cys Arg Phe Arg
     30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 409:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 138 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.3
seq ALSLLLVSGSLLP/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 409:

```
Met Ser Gly Ser Ser Leu Pro Ser Ala Leu Ala Leu Ser Leu Leu Leu
            -20                 -15                 -10

Val Ser Gly Ser Leu Leu Pro Gly Pro Gly Ala Ala Gln Asn Glu Pro
```

```
        -5                  1                 5

Arg Ile Val Thr Ser Glu Glu Val Ile Ile Arg Asp Ser Pro Val Leu
10                  15                  20                  25

Pro Val Thr Leu Gln Cys Asn Leu Thr Ser Ser Ser His Thr Leu Thr
                30                  35                  40

Tyr Ser Tyr Trp Thr Lys Asn Gly Val Glu Leu Ser Ala Thr Arg Lys
            45                  50                  55

Asn Ala Ser Asn Met Glu Tyr Arg Ile Asn Lys Pro Arg Ala Glu Asp
        60                  65                  70

Ser Gly Glu Tyr His Cys Val Tyr His Phe Val Ser Ala Pro Lys Ala
    75                  80                  85

Asn Ala Thr Ile Glu Val Lys Ala Ala Pro Asp Ile Thr Gly His Lys
90                  95                  100                 105

Arg Ser Xaa Asn Lys Asn Glu Gly Gln Asp
                110                 115
```

(2) INFORMATION FOR SEQ ID NO: 410:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 96 amino acids
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: -36..-1
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 7.2
  seq IMLLSLAAFSVIS/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 410:

```
Met Ala Val His Asp Leu Ile Phe Trp Arg Asp Val Lys Lys Thr Gly
    -35                 -30                 -25

Phe Val Phe Gly Thr Thr Leu Ile Met Leu Leu Ser Leu Ala Ala Phe
-20                 -15                 -10                 -5

Ser Val Ile Ser Val Val Ser Tyr Leu Ile Leu Ala Leu Leu Ser Val
                1                   5                   10

Thr Ile Ser Phe Arg Ile Tyr Lys Ser Val Ile Gln Ala Val Gln Lys
        15                  20                  25

Ser Glu Glu Gly His Pro Phe Lys Ala Tyr Leu Asp Val Asp Ile Thr
```

30 35 40

Leu Ser Ser Glu Ala Phe His Asn Tyr Met Asn Ala Ala Met Val His
45 50 55 60

(2) INFORMATION FOR SEQ ID NO: 411:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq LLWTLLLFAAPFG/LL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 411:

Met Xaa Gly Ser Val Glu Cys Thr Xaa Gly Trp Gly His Cys Ala Pro
-30 -25 -20

Ser Pro Leu Leu Leu Trp Thr Leu Leu Leu Phe Ala Ala Pro Phe Gly
-15 -10 -5

Leu Leu Gly Glu Lys Thr Arg Gln Val Ser Leu Glu Val Ile Pro Asn
1 5 10 15

Trp Leu Gly Pro Leu Gln Asn Leu Leu His Ile Arg Ala Val Gly Thr
20 25 30

Asn Ser Thr Leu His Tyr Val Trp Ser Ser Leu Gly Pro Leu Ala Val
35 40 45

Val Met Val Ala Thr Asn Thr Pro Pro Gly
50 55

(2) INFORMATION FOR SEQ ID NO: 412:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq LIFLCGAALLAVG/IW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 412:

```
Met Gln Cys Phe Ser Phe Ile Lys Thr Met Met Ile Leu Phe Asn Leu
                -25                 -20                 -15

Leu Ile Phe Leu Cys Gly Ala Ala Leu Leu Ala Val Gly Ile Trp Val
            -10                 -5                  1

Ser Ile Asp Gly Ala Ser Phe Leu Lys Ile Phe Gly Pro Leu Ser Ser
    5                   10                  15

Ser Ala Met Gln Phe Val Asn Val Gly Tyr Phe Leu Ile Ala Ala Gly
20                  25                  30                  35

Val Val Val Phe Ala Leu Gly Phe Leu Gly Cys Xaa Gly Ala Lys Xaa
                40                  45                  50

Glu Xaa Lys Cys Ala Leu Val Thr Phe Phe Phe Ile Leu Leu Leu Ile
            55                  60                  65

Phe
```

(2) INFORMATION FOR SEQ ID NO: 413:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.1
seq LLWTLLLFAAPFG/LL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 413:

```
Met Arg Gly Ser Val Glu Cys Thr Trp Gly Xaa Gly His Cys Ala Pro
        -30                 -25                 -20

Ser Pro Leu Leu Leu Trp Thr Leu Leu Leu Phe Ala Ala Pro Phe Gly
```

-15 -10 -5

Leu Leu Gly Glu Lys Thr His Gln Val Ser Leu Glu Val Ile Pro Asn
1 5 10 15

Trp Leu Gly Pro Leu Gln Asn Leu Leu His Ile Arg Xaa Val Gly Thr
20 25 30

Asn Ser Thr Leu His Tyr Val Trp Ser Ser Leu Gly Pro Leu Ala Val
35 40 45

Val Met Val Ala Thr Asn Thr Pro His Ser Thr Leu Ser Val Asn Trp
50 55 60

Ser Leu Leu Leu Ser Pro Glu Pro Asp Gly Gly Leu Met Val Leu Pro
65 70 75 80

Lys Asp Ser Ile Gln Phe Ser Ser
85

(2) INFORMATION FOR SEQ ID NO: 414:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymphocytes

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7
seq LRLLKLAATSASA/RV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 414:

Met Ala Leu Arg Leu Leu Lys Leu Ala Ala Thr Ser Ala Ser Ala Arg
-15 -10 -5 1

Val Val Ala Ala Gly Ala Gln Arg Val Arg Gly Ile His Ser Ser Val
5 10 15

Gln Cys Lys Leu Arg Tyr Gly Met Trp His Phe Leu Leu Gly Asp Lys
20 25 30

Ala Ser Lys Arg Leu Thr Glu Arg Ser Arg Val Ile Thr Val Asp Gly
35 40 45

Asn Met
50

(2) INFORMATION FOR SEQ ID NO: 415:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -65..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7
seq IGHFLCLVILVYC/AE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 415:

```
Met Pro Ser Ala Phe Ser Val Ser Ser Phe Pro Val Ser Ile Pro Ala
-65                 -60                 -55                 -50

Val Leu Thr Gln Thr Asp Trp Thr Glu Pro Trp Leu Met Gly Leu Ala
                -45                 -40                 -35

Thr Phe His Ala Leu Cys Val Leu Leu Thr Cys Leu Ser Ser Arg Ser
            -30                 -25                 -20

Tyr Arg Leu Gln Ile Gly His Phe Leu Cys Leu Val Ile Leu Val Tyr
        -15                 -10                  -5

Cys Ala Glu Tyr Ile Asn Glu Ala Ala Ala Met Asn Trp Arg Leu Phe
      1               5                  10                  15

Ser Lys Tyr Gln Tyr Phe Asp Ser Arg Gly Met Phe Ile Ser Ile Val
                 20                  25                  30

Phe Ser Ala Pro Leu Leu Val Asn Ala Met Ile Ile Val Val Met Trp
             35                  40                  45

Val Trp Lys Thr
         50
```

(2) INFORMATION FOR SEQ ID NO: 416:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 163 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -154..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.7
seq ALGILVVAGCSFA/IR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 416:

```
Met Ala Leu Pro His Gln Glu Pro Lys Pro Gly Asp Leu Ile Glu Ile
                -150                -145                -140

Phe Arg Leu Gly Tyr Glu His Trp Ala Leu Tyr Ile Xaa Asp Gly Tyr
            -135                -130                -125

Val Ile His Leu Ala Pro Pro Ser Glu Tyr Pro Gly Ala Gly Ser Ser
        -120                -115                -110

Ser Val Phe Ser Val Leu Ser Asn Ser Ala Glu Val Lys Arg Glu Arg
    -105                -100                -95

Leu Glu Asp Val Val Gly Gly Cys Cys Tyr Arg Val Asn Asn Ser Leu
-90                 -85                 -80                 -75

Asp His Glu Tyr Gln Pro Arg Pro Val Glu Val Ile Ile Ser Ser Ala
                -70                 -65                 -60

Lys Glu Met Val Gly Gln Lys Met Lys Tyr Ser Ile Val Ser Arg Asn
            -55                 -50                 -45

Cys Glu His Phe Val Thr Gln Leu Arg Tyr Gly Lys Ser Arg Cys Lys
        -40                 -35                 -30

Gln Val Glu Lys Ala Lys Val Glu Val Gly Val Ala Thr Ala Leu Gly
    -25                 -20                 -15

Ile Leu Val Val Ala Gly Cys Ser Phe Ala Ile Arg Arg Tyr Gln Lys
-10                 -5                  1                   5

Lys Ala Thr
```

(2) INFORMATION FOR SEQ ID NO: 417:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: -70..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.7
seq LAFSLPALPLAEL/QP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 417:

Met Ala Ala Ser Thr Ser Met Val Pro Val Ala Val Thr Ala Ala Val
-70 -65 -60 -55

Ala Pro Val Leu Ser Ile Asn Ser Asp Phe Ser Asp Leu Arg Glu Ile
-50 -45 -40

Lys Lys Gln Leu Leu Leu Ile Ala Gly Leu Thr Arg Glu Arg Gly Leu
-35 -30 -25

Leu His Ser Ser Lys Trp Ser Ala Glu Leu Ala Phe Ser Leu Pro Ala
-20 -15 -10

Leu Pro Leu Ala Glu Leu Gln Pro Pro Pro Pro Ile Thr Glu Glu Asp
-5 1 5 10

Ala Gln Asp Met Asp Ala Tyr Thr Leu Ala Lys Ala Tyr Phe Asp Val
15 20 25

Lys Glu Tyr Asp Arg Ala Ala His Phe Leu His Gly Cys Asn Ala Arg
30 35 40

Xaa Ala Tyr Phe Leu Tyr Met Tyr Ser Arg Tyr Leu Ser
45 50 55

(2) INFORMATION FOR SEQ ID NO: 418:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.6
seq KMVHLLVLSGAWG/MQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 418:

Met Glu Glu Gly Gly Asn Leu Gly Gly Leu Ile Lys Met Val His Leu
-20 -15 -10

Leu Val Leu Ser Gly Ala Trp Gly Met Gln Met Trp Val Thr Phe Val

-5 1 5

Ser Gly Phe Leu Leu Phe Arg Ser Leu Pro Arg His Thr Phe Gly Leu
10 15 20

Val Gln Ser Lys Leu Phe Pro Phe Tyr Phe His Ile Ser Met Gly Cys
25 30 35 40

Ala Phe Ile Asn Xaa Cys Ile Leu Ala Ser Gln His Ala Trp Ala Gln
45 50 55

Leu Thr Phe Trp Glu Ala Ser Gln Leu Tyr Leu Leu Phe Leu
60 65 70

(2) INFORMATION FOR SEQ ID NO: 419:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 87 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -81..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.6
seq LLLASGTTLFCTS/FY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 419:

Met Ala Gly Pro Ala Ala Ala Phe Arg Arg Leu Gly Ala Leu Ser Gly
-80 -75 -70

Ala Ala Ala Leu Gly Phe Ala Ser Tyr Gly Ala His Gly Ala Xaa Phe
-65 -60 -55 -50

Pro Asp Ala Tyr Gly Lys Glu Leu Phe Asp Lys Ala Asn Lys His His
-45 -40 -35

Phe Leu His Ser Leu Ala Leu Leu Gly Val Pro His Cys Arg Lys Pro
-30 -25 -20

Leu Trp Ala Gly Leu Leu Leu Ala Ser Gly Thr Thr Leu Phe Cys Thr
-15 -10 -5

Ser Phe Tyr Tyr Gln Ala Gln
1 5

(2) INFORMATION FOR SEQ ID NO: 420:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq LLTLLLPPPPLYT/RH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 420:

```
Met Gly His Arg Phe Leu Arg Gly Leu Leu Thr Leu Leu Leu Pro Pro
    -20                 -15                 -10

Pro Pro Leu Tyr Thr Arg His Arg Met Leu Gly Pro Glu Ser Val Pro
 -5                  1               5                  10

Pro Pro Lys Arg Ser Arg Ser Lys Leu Met Ala Pro Pro Arg
             15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO: 421:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -20..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.5
seq ILFLLPSICSSNS/TG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 421:

```
Met Glu Leu Leu Gln Val Thr Ile Leu Phe Leu Leu Pro Ser Ile Cys
-20                 -15                 -10                  -5

Ser Ser Asn Ser Thr Gly Val Leu Glu Ala Ala Asn Asn Ser Leu Val
```

1 5 10

Val Thr Thr Thr Xaa Pro Ser Ile Thr Thr Pro Asn Thr Glu Ser Leu
15 20 25

Gln Lys Asn Val Val Thr Pro Thr Thr Gly Thr Thr Xaa Lys Gly Thr
30 35 40

Ile Thr Asn Glu Leu Leu Lys Met Ser Leu Met Ser Thr Ala Xaa Phe
45 50 55 60

(2) INFORMATION FOR SEQ ID NO: 422:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq VLMRLVASAYSIA/QK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 422:

Met Ala Ser Ser Asn Thr Val Leu Met Arg Leu Val Ala Ser Ala Tyr
-15 -10 -5

Ser Ile Ala Gln Lys Ala Gly Met Ile Val Arg Arg Val Ile Ala Glu
1 5 10

Gly Asp Leu Gly Ile Val Glu Lys Thr Cys Ala Thr Asp Leu Gln Thr
15 20 25

Lys Ala Asp Arg Leu Ala Gln Met Ser Ile Cys Ser Ser Leu Ala Arg
30 35 40 45

Lys Phe Pro Lys Leu Thr Ile Ile Gly Glu Glu Asp Leu Pro Ser Glu
50 55 60

Glu Val Asp Gln Glu Leu Ile Glu Asp Ser Gln Trp Glu Glu Ile Leu
65 70 75

Lys Gln Pro Cys Pro Ser Gln Tyr Ser Ala Ile Lys Glu Glu Asp Leu
80 85 90

Val Val Trp Val Asp
95

(2) INFORMATION FOR SEQ ID NO: 423:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq SSCVLLTALVALA/AY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 423:

```
Met Arg Ser Ser Cys Val Leu Leu Thr Ala Leu Val Ala Leu Ala Ala
-15                 -10                 -5                   1

Tyr Tyr Val Tyr Ile Pro Leu Pro Gly Ser Val Ser Asp Pro Trp Lys
            5                   10                  15

Leu Met Leu Leu Asp Ala Thr Phe Arg Gly Ala Gln Gln Val Ser Asn
        20                  25                  30

Leu Ile His Tyr Leu Gly Leu Ser His His Leu Leu Ala Leu Asn Phe
    35                  40                  45

Ile Ile Val Ser Phe Gly Lys Lys Ser Ala Trp Ser Ser Ala Gln Val
 50                  55                  60                  65

Lys Val Thr Asp Thr Asp Phe Asp Gly Val Glu Val Arg Val Phe Glu
                70                  75                  80

Gly Pro Pro Lys Pro Glu Glu Pro Leu Lys Arg Ser Val Val Tyr Ile
            85                  90                  95

His Gly Xaa Gly Trp
        100
```

(2) INFORMATION FOR SEQ ID NO: 424:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens

(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq GVGLVTLLGLAVG/SY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 424:

```
Met Gly Ile Gln Thr Ser Pro Val Leu Leu Ala Ser Leu Gly Val Gly
    -25                 -20                 -15

Leu Val Thr Leu Leu Gly Leu Ala Val Gly Ser Tyr Leu Val Arg Arg
-10                 -5                  1               5

Ser Arg Arg Pro Gln Val Thr Leu Leu Asp Pro Asn Glu Lys Tyr Leu
            10                  15                  20

Leu Arg Leu Leu Asp Lys Thr Thr Val Ser His Asn Thr Lys Arg Phe
        25                  30                  35

Arg Phe Ala Leu Pro Thr Ala His His Met
     40                  45
```

(2) INFORMATION FOR SEQ ID NO: 425:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -69..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.3
seq ILLIVLFLDAVRE/VR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 425:

```
Met Thr Leu Gln Trp Ala Ala Val Ala Thr Phe Leu Tyr Ala Glu Ile
                -65                 -60                 -55

Gly Leu Ile Leu Ile Phe Cys Leu Pro Phe Ile Pro Pro Gln Arg Trp
            -50                 -45                 -40

Gln Lys Ile Phe Ser Phe Asn Val Trp Gly Lys Ile Ala Thr Phe Trp
```

```
        -35                 -30                 -25

Asn Lys Ala Phe Leu Thr Ile Ile Ile Leu Leu Ile Val Leu Phe Leu
    -20                 -15                 -10

Asp Ala Val Arg Glu Val Arg Lys Tyr Ser Ser Val His Thr Ile Glu
 -5                  1                   5                  10

Lys Ser Ser Thr Ser Arg Pro Arg
             15
```

(2) INFORMATION FOR SEQ ID NO: 426:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 94 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -85..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.2
seq FLDFCVYIPLSWG/FC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 426:

```
Met Pro Ser Glu Gly Arg Cys Trp Glu Thr Leu Lys Ala Leu Arg Ser
-85                 -80                 -75                 -70

Ser Asp Lys Gly Arg Leu Cys Tyr Tyr Arg Asp Trp Leu Leu Arg Arg
                -65                 -60                 -55

Glu Val Ser Gly Gly Pro Gly Gly Arg Arg Pro Phe Arg Pro Leu Ala
            -50                 -45                 -40

Thr Glu Thr Phe Ser Leu Ala Val Gly Thr Phe Cys Ser Arg Glu Pro
        -35                 -30                 -25

Val Gln Ser Asn Asn Leu His Leu Phe Leu Asp Phe Cys Val Tyr Ile
    -20                 -15                 -10

Pro Leu Ser Trp Gly Phe Cys Pro Leu Gln Pro Ile Leu Ala
 -5                  1                   5
```

(2) INFORMATION FOR SEQ ID NO: 427:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 113 amino acids

(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.2
seq AILGSTWVALTTG/AL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 427:

```
Met Thr Lys Leu Ala Gln Trp Leu Trp Gly Leu Ala Ile Leu Gly Ser
                -20                 -15                 -10

Thr Trp Val Ala Leu Thr Thr Gly Ala Leu Gly Leu Glu Leu Pro Leu
             -5                   1               5

Ser Cys Gln Glu Val Leu Trp Pro Leu Pro Ala Tyr Leu Leu Val Ser
     10                  15                  20

Ala Gly Cys Tyr Ala Leu Gly Thr Val Gly Tyr Arg Val Ala Thr Phe
 25                  30                  35                  40

His Asp Cys Glu Asp Ala Ala Arg Glu Leu Gln Ser Gln Ile Gln Glu
                 45                  50                  55

Ala Arg Ala Asp Leu Ala Arg Xaa Gly Cys Ala Ser Asp Ser Leu Xaa
             60                  65                  70

Pro Phe Leu Cys Gly Gln Pro Phe Leu Pro Phe Pro Ile Lys Glu Pro
         75                  80                  85

Gly
```

(2) INFORMATION FOR SEQ ID NO: 428:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 6.2
seq FLVSNMLLAEAYG/SG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 428:

```
Met Leu Leu Ala Trp Val Gln Ala Phe Leu Val Ser Asn Met Leu Leu
    -20                 -15                 -10

Ala Glu Ala Tyr Gly Ser Gly Gly Cys Phe Trp Asp Asn Gly His Leu
 -5                  1               5                  10

Tyr Arg Glu Asp Gln Thr Ser Pro Ala Pro Gly Leu Arg Cys Leu Asn
            15                  20                  25

Trp Leu Asp Ala Gln Ser Gly
        30
```

(2) INFORMATION FOR SEQ ID NO: 429:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 126 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -41..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.2
seq SVLVLLLLAVLYE/GI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 429:

```
Met Ala Met His Phe Ile Phe Ser Asp Thr Ala Val Leu Leu Phe His
    -40                 -35                 -30

Phe Trp Ser Val His Ser Pro Ala Gly Met Ala Leu Ser Val Leu Val
-25                 -20                 -15                 -10

Leu Leu Leu Leu Ala Val Leu Tyr Glu Gly Ile Lys Val Gly Lys Ala
                 -5                  1               5

Lys Leu Leu Asn Gln Val Leu Val Asn Leu Pro Thr Ser Ile Ser Gln
         10                  15                  20

Gln Thr Ile Ala Glu Thr Asp Gly Asp Ser Ala Gly Ser Asp Ser Phe
     25                  30                  35

Pro Val Gly Arg Thr His His Arg Trp Tyr Leu Cys His Phe Gly Gln
 40                  45                  50                  55
```

```
Ser Leu Ile His Val Ile Gln Val Val Ile Gly Tyr Phe Ile Met Leu
                60                  65                  70

Ala Val Met Ser Tyr Asn Thr Trp Ile Phe Leu Gly Val Val
            75                  80                  85
```

(2) INFORMATION FOR SEQ ID NO: 430:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -75..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.2
seq VVXXSVLXTTCXS/SQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 430:

```
Met Lys Gln Val His Gln Cys Ile Glu Arg Cys His Val Pro Leu Ala
-75                 -70                 -65                 -60

Gln Ala Gln Ala Leu Val Thr Ser Glu Leu Glu Lys Phe Gln Asp Arg
                -55                 -50                 -45

Leu Ala Arg Cys Thr Met His Cys Asn Asp Lys Ala Lys Asp Ser Ile
            -40                 -35                 -30

Asp Ala Gly Xaa Lys Glu Leu Gln Val Lys Gln Gln Leu Xaa Val Val
        -25                 -20                 -15

Xaa Xaa Ser Val Leu Xaa Thr Thr Cys Xaa Ser Ser Gln Leu
    -10                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 431:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.2
seq LLAALMLVAMLQL/LY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 431:

```
Met Gln Met Ser Tyr Ala Ile Arg Cys Ala Phe Tyr Gln Leu Leu Leu
        -25                 -20                 -15

Ala Ala Leu Met Leu Val Ala Met Leu Gln Leu Leu Tyr Leu Ser Leu
    -10                 -5                  1                   5

Leu Ser Gly Leu His Gly Pro
                10
```

(2) INFORMATION FOR SEQ ID NO: 432:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 60 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.1
seq IILLIHTMQVCTT/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 432:

```
Met Met Thr Gln Thr Cys Ile Ile Leu Leu Ile His Thr Met Gln Val
                -15                 -10                 -5

Cys Thr Thr His Pro Thr Val Leu Ser His Thr Leu Leu Gln Arg Pro
            1                   5                   10

Lys Pro Thr Asp Leu Phe Pro Lys Ala Thr Pro Thr Thr Ala Pro Met
    15                  20                  25

Pro Leu Arg Met Arg Pro Pro Gln Cys Leu Pro Glu
30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 433:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.1
seq LFLTCLFWPLAAL/NV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 433:

```
Met Xaa Xaa His Leu Gln Thr Arg Pro Leu Phe Leu Thr Cys Leu Phe
        -20                 -15                 -10

Trp Pro Leu Ala Ala Leu Asn Val Asn Ser Thr Phe Glu Cys Leu Ile
     -5                  1               5                  10

Leu Gln Cys Ser Val Gly Ile
                15
```

(2) INFORMATION FOR SEQ ID NO: 434:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 92 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.1
seq LMAFLLSFYLIFT/NE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 434:

```
Met Ala Ala Asn Tyr Ser Ser Thr Xaa Thr Arg Arg Glu His Val Lys
        -50                 -45                 -40

Val Lys Thr Ser Ser Gln Pro Gly Phe Leu Glu Arg Leu Ser Glu Thr
    -35                 -30                 -25
```

```
Ser Gly Gly Met Phe Val Gly Leu Met Ala Phe Leu Leu Ser Phe Tyr
-20                 -15                 -10                  -5

Leu Ile Phe Thr Asn Glu Gly Arg Ala Leu Lys Thr Ala Thr Ser Leu
                  1               5                  10

Ala Glu Gly Leu Ser Leu Val Val Ser Pro Asp Ser Ile His Ser Val
        15                  20                  25

Ala Pro Glu Asn Glu Gly Xaa Leu Val His Ile Ile
    30                  35                  40
```

(2) INFORMATION FOR SEQ ID NO: 435:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.1
seq LEMLTAFASHIRA/RD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 435:

```
Met Arg Gly Ala His Leu Thr Ala Leu Glu Met Leu Thr Ala Phe Ala
    -20                 -15                 -10

Ser His Ile Arg Ala Arg Asp Ala Ser Gly
 -5                      1               5
```

(2) INFORMATION FOR SEQ ID NO: 436:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6.1
seq IILLIHTMQVCTT/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 436:

```
Met Val His Lys Pro Met Met Thr Gln Thr Cys Ile Ile Leu Leu Ile
                -20                 -15                 -10

His Thr Met Gln Val Cys Thr Thr His Pro Thr Val Leu Ser His Thr
            -5                   1               5

Leu Leu Gln Arg Pro Lys Pro Thr Asp Leu Phe Pro Lys Ala Thr Pro
     10                  15                  20

Thr Thr
 25
```

(2) INFORMATION FOR SEQ ID NO: 437:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -28..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6
seq IGLMFLMLGCALP/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 437:

```
Met Ala Gly Ile Lys Ala Leu Ile Ser Leu Ser Phe Gly Gly Ala Ile
            -25                 -20                 -15

Gly Leu Met Phe Leu Met Leu Gly Cys Ala Leu Pro Ile Tyr Asn Lys
        -10                 -5                   1

Tyr Trp Pro Thr
  5
```

(2) INFORMATION FOR SEQ ID NO: 438:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 6
seq LLFPLTLVRSFWS/DM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 438:

```
Met Ser Leu Met Pro Lys Met His Leu Leu Phe Pro Leu Thr Leu Val
    -20                 -15                 -10

Arg Ser Phe Trp Ser Asp Met Met Asp Ser Ala Gln Ser Phe Ile Thr
 -5                  1               5                  10

Ser Ser Trp Thr Phe Tyr Leu Gln Ala Asp Asp Gly Lys Ile Val Ile
            15                  20                  25

Phe Gln Ser Lys Pro Glu Ile Gln Tyr Ala Pro His Leu Glu Gln Glu
        30                  35                  40

Pro Thr Asn Leu Arg Glu Ser Ser Leu Ser Lys Met Ser Tyr Leu Gln
    45                  50                  55

Met Arg Asn Ser Gln Ala His Arg
 60                 65
```

(2) INFORMATION FOR SEQ ID NO: 439:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 115 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -87..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.9
seq SNILLASVGSVLG/AC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 439:

```
Met Met Lys Arg Ala Ala Ala Ala Ala Val Gly Gly Ala Leu Ala Val
```

-85 -80 -75

Gly Ala Val Pro Val Val Leu Ser Ala Met Gly Phe Thr Gly Ala Gly
-70 -65 -60

Ile Ala Ala Ser Ser Ile Ala Ala Lys Met Met Ser Ala Ala Ala Ile
-55 -50 -45 -40

Ala Asn Gly Gly Gly Val Ser Ala Gly Ser Leu Val Ala Thr Leu Gln
-35 -30 -25

Ser Val Gly Ala Ala Gly Leu Ser Thr Ser Ser Asn Ile Leu Leu Ala
-20 -15 -10

Ser Val Gly Ser Val Leu Gly Ala Cys Leu Gly Asn Ser Pro Ser Xaa
-5 1 5

Ser Leu Pro Ala Glu Pro Xaa Xaa Xaa Glu Asp Glu Ala Arg Glu Asn
10 15 20 25

Val Pro Pro

(2) INFORMATION FOR SEQ ID NO: 440:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.8
seq VTIILLLSCXFWA/VK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 440:

Met Val Thr Ile Ile Leu Leu Leu Ser Cys Xaa Phe Trp Ala Val Lys
-10 -5 1

Asn Val Thr Xaa Arg Xaa Met Val Gly Leu Arg Trp Trp Asn His Ile
5 10 15

(2) INFORMATION FOR SEQ ID NO: 441:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 130 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -87..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.8
seq SNILLASVGSVSG/AC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 441:

```
Met Xaa Lys Arg Ala Ala Ala Ala Ala Val Gly Gly Ala Leu Ala Val
        -85                 -80                 -75

Gly Ala Val Pro Val Val Leu Ser Ala Met Gly Phe Thr Gly Ala Gly
    -70                 -65                 -60

Ile Ala Ala Ser Ser Ile Ala Ala Lys Met Met Ser Ala Ala Ala Ile
-55                 -50                 -45                 -40

Ala Asn Gly Gly Gly Val Ser Ala Gly Ser Leu Val Ala Thr Leu Gln
                -35                 -30                 -25

Ser Val Gly Ala Ala Gly Leu Ser Thr Ser Ser Asn Ile Leu Leu Ala
            -20                 -15                 -10

Ser Val Gly Ser Val Ser Gly Ala Cys Leu Gly Asn Ser Pro Ser Ser
        -5                   1               5

Ser Leu Pro Ala Glu Pro Glu Ala Lys Glu Asp Glu Ala Arg Glu Asn
 10                  15                  20                  25

Val Pro Gln Gly Glu Pro Pro Lys Pro Pro Leu Lys Ser Glu Lys His
                 30                  35                  40

Glu Arg
```

(2) INFORMATION FOR SEQ ID NO: 442:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -89..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq DLSLLSLPPGTSP/VG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 442:

Met Ser Gln Asp Gly Gly Xaa Gly Glu Leu Lys His Met Val Met Ser
-85 -80 -75

Phe Arg Val Ser Glu Leu Gln Val Leu Leu Gly Xaa Xaa Gly Arg Asn
-70 -65 -60

Lys Ser Gly Arg Lys His Glu Leu Leu Ala Lys Ala Leu His Leu Leu
-55 -50 -45

Lys Ser Ser Cys Ala Pro Ser Val Gln Met Lys Ile Lys Glu Leu Tyr
-40 -35 -30

Arg Arg Arg Phe Pro Arg Lys Thr Leu Gly Pro Ser Asp Leu Ser Leu
-25 -20 -15 -10

Leu Ser Leu Pro Pro Gly Thr Ser Pro Val Gly Ser Pro Gly Pro Leu
-5 1 5

Ala Pro Ile Pro Pro Thr Xaa Leu Ala Xaa Ala Xaa Cys Trp Ala Pro
10 15 20

Ser Val Arg Trp Thr Cys
25

(2) INFORMATION FOR SEQ ID NO: 443:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 46 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq LLLPRVLLTMASG/SL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 443:

Met Pro Xaa Leu Leu Pro Val Ala Ser Arg Leu Leu Leu Leu Pro Arg
-20 -15 -10

Val Leu Leu Thr Met Ala Ser Gly Ser Leu Arg Xaa Ser Xaa Arg Arg
-5 1 5

```
Pro Arg Ile Pro Xaa Leu Ala Thr Phe Arg Xaa Arg Ser Leu
     10                 15                  20
```

(2) INFORMATION FOR SEQ ID NO: 444:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq IFSFLDIVTLCRC/AQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 444:

```
Met Val Phe Ser Asn Asn Asp Glu Gly Leu Ile Asn Lys Lys Leu Pro
-35                 -30                 -25                 -20

Lys Glu Leu Leu Leu Arg Ile Phe Ser Phe Leu Asp Ile Val Thr Leu
                -15                 -10                  -5

Cys Arg Cys Ala Gln Ile Xaa Lys Ala Trp Asn Ile Leu Ala Leu Asp
              1               5                  10

Gly Ser Asn Trp Gln Arg Ile Asp Leu Phe Asn Phe Gln Thr Asp Val
     15                  20                  25

Glu Gly Arg Val Val Glu Asn Ile Ser Lys Arg Cys Gly Gly Phe Leu
 30                  35                  40                  45

Arg Lys Leu Ser Leu Arg Gly Cys Ile Gly Val Gly Xaa Ser Ser Leu
                 50                  55                  60

Xaa Thr Phe Ala Gln Asn Cys Arg Asn Ile Glu His Leu Asn Leu Asn
             65                  70                  75

Gly Cys Thr Lys Ile Thr Xaa Ser Thr Cys
         80                  85
```

(2) INFORMATION FOR SEQ ID NO: 445:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.7
seq IFSFLDIVTLCRC/AQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 445:

```
Met Val Phe Ser Asn Asn Asp Glu Gly Leu Ile Asn Lys Lys Leu Pro
-35                 -30                 -25                 -20

Lys Glu Leu Leu Leu Arg Ile Phe Ser Phe Leu Asp Ile Val Thr Leu
                -15                 -10                  -5

Cys Arg Cys Ala Gln Ile Ser Lys Ala Trp Asn Ile Leu Ala Leu Asp
             1                   5                  10

Gly Ser Asn Trp Gln Gly
    15
```

(2) INFORMATION FOR SEQ ID NO: 446:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 145 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -112..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq SSCILPWLSKTNS/CP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 446:

```
Met Ala Ser Tyr Phe Asp Glu His Asp Cys Glu Pro Ser Asp Pro Glu
    -110                -105                -100

Gln Glu Thr Arg Thr Asn Met Leu Leu Glu Leu Ala Arg Ser Leu Phe
    -95                 -90                 -85
```

Asn Arg Met Asp Phe Glu Asp Leu Gly Leu Val Val Asp Trp Asp His
-80 -75 -70 -65

His Leu Pro Pro Pro Ala Ala Lys Thr Val Val Glu Asn Leu Pro Arg
-60 -55 -50

Thr Val Ile Arg Gly Ser Gln Ala Glu Leu Lys Cys Pro Val Cys Leu
-45 -40 -35

Leu Glu Phe Glu Glu Glu Glu Thr Ala Ile Glu Met Pro Cys His His
-30 -25 -20

Leu Phe His Ser Ser Cys Ile Leu Pro Trp Leu Ser Lys Thr Asn Ser
-15 -10 -5

Cys Pro Leu Cys Arg Tyr Glu Leu Pro Thr Asp Asp Asp Thr Tyr Glu
1 5 10 15

Glu His Arg Arg Asp Lys Ala Arg Lys Gln Gln Gln Gln His Arg Pro
20 25 30

Xaa

(2) INFORMATION FOR SEQ ID NO: 447:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LILSLQVCRPATL/DQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 447:

Met Pro Leu Ile Leu Ser Leu Gln Val Cys Arg Pro Ala Thr Leu Asp
-15 -10 -5 1

Gln Ala Thr Arg Ala Thr Thr Pro Cys Arg Leu Ser Gln Gly Cys Gln
5 10 15

Gln His Pro Thr Gln Cys Ser Thr His His Leu Thr Gln Pro Ser Pro
20 25 30

Trp Ala His Arg Xaa Thr Thr Arg Pro Trp Leu Glu Glu Gln Pro Arg
35 40 45

Pro Gly

(2) INFORMATION FOR SEQ ID NO: 448:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -73..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LRRLLGCLTLTLS/GR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 448:

```
Met Leu Gly Ile Thr Ser Cys Ser Asp Gln Gln Ala Lys Glu Gly Glu
            -70                 -65                 -60

Gly Leu Glu Gly Ser Ser Thr Gly Ser Ser Ser Gly Asn His Gly Gly
        -55                 -50                 -45

Ser Gly Gly Gly Asn Gly His Lys Pro Gly Cys Glu Lys Pro Gly Asn
    -40                 -35                 -30

Glu Ala Arg Gly Ser Gly Asn Leu Gly Phe Arg Thr Leu Arg Arg Leu
-25                 -20                 -15                 -10

Leu Gly Cys Leu Thr Leu Thr Leu Ser Gly Arg Ile
                -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 449:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 5.6
seq ALKLASWTSMALA/AS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 449:

```
Met Ala Arg Lys Ala Leu Lys Leu Ala Ser Trp Thr Ser Met Ala Leu
        -15                 -10                 -5

Ala Ala Ser Gly Ile Tyr Phe Tyr Ser Asn Lys Tyr Leu Asp Pro Asn
     1              5                  10                  15

Asp Phe Gly Ala Val Arg Val Gly Arg Ala Val Ala Thr Thr Ala Val
                20                  25                  30

Ile Ser Xaa Asp Tyr Leu Thr Ser Leu Lys Ser Val Pro Tyr Gly Ser
            35                  40                  45

Glu Glu Tyr Leu Gln Leu Arg Ser Lys Val His Leu Arg Ser Ala Arg
        50                  55                  60

Arg Leu Cys Xaa Xaa Cys Cys Ala Asn Arg Gly
    65                  70
```

(2) INFORMATION FOR SEQ ID NO: 450:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq AALPAWLSLQSRA/RS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 450:

```
Met Ala Ala Ala Ala Leu Pro Ala Trp Leu Ser Leu Gln Ser Arg Ala
    -15                 -10                 -5

Arg Ser Leu Arg Ala Phe Ser Thr Ala Val Tyr Ser Ala Thr Pro Val
  1               5                  10                  15

Pro Thr Pro Ser Leu Pro Glu Arg Thr Pro Gly Asn Glu Arg Pro Pro
            20                  25                  30

Xaa Arg Lys Ala Leu Pro Pro Arg Thr Glu Lys Met Ala Val Asp Gln
        35                  40                  45
```

Asp Trp Pro Ser Val Tyr Pro Val Ala Ala Pro Xaa Lys Pro Ser Ala
50 55 60

Val Pro Leu Pro Val Arg Met Gly Tyr Pro Val Lys Lys Gly Val Pro
65 70 75 80

Met Ala Lys Glu Gly Asn Leu Glu Leu Leu Lys Ile Pro Asn Phe Leu
85 90 95

His Leu Thr Pro Val Ala Ile Lys Lys His Cys Xaa Ala Leu Lys Asp
100 105 110

Phe Cys Thr Glu
115

(2) INFORMATION FOR SEQ ID NO: 451:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 112 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -65..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq CMLTLXXLSFILA/GL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 451:

Met Val Lys Ile Ala Phe Asn Thr Pro Thr Ala Val Gln Lys Glu Glu
-65 -60 -55 -50

Ala Arg Gln Asp Val Glu Ala Leu Leu Ser Arg Thr Val Arg Thr Gln
-45 -40 -35

Ile Leu Thr Gly Lys Glu Leu Arg Val Ala Thr Gln Glu Lys Glu Gly
-30 -25 -20

Ser Ser Gly Arg Cys Met Leu Thr Leu Xaa Xaa Leu Ser Phe Ile Leu
-15 -10 -5

Ala Gly Leu Ile Val Gly Gly Ala Cys Ile Tyr Lys Tyr Phe Met Pro
1 5 10 15

Lys Ser Thr Ile Tyr Arg Gly Xaa Met Cys Phe Phe Asp Ser Glu Asp
20 25 30

Pro Ala Asn Ser Leu Arg Gly Gly Glu Pro Asn Phe Leu Pro Val Thr
35 40 45

(2) INFORMATION FOR SEQ ID NO: 452:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 59 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Thyroid

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -48..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LLLSFVWMPALLP/DG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 452:

```
Met Ile Gly Ser Gly Leu Ala Gly Ser Gly Gly Ala Gly Gly Pro Ser
            -45                 -40                 -35

Ser Thr Val Thr Trp Cys Ala Leu Xaa Ser Asn His Val Ala Ala Thr
        -30                 -25                 -20

Gln Ala Ser Leu Leu Leu Ser Phe Val Trp Met Pro Ala Leu Leu Pro
    -15                 -10                  -5

Asp Gly Leu Pro Pro Phe Val Ala Thr Pro Met
  1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 453:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LXGFLFXVIVLTS/WI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 453:

Met Ser Gly Ala Gln Leu Xaa Gly Phe Leu Phe Xaa Val Ile Val Leu
-15 -10 -5

Thr Ser Trp Ile Thr Ile Phe Gln Ile Tyr Arg Pro Arg Trp Gly Cys
1 5 10

Pro Trp Gly Leu Pro Leu Leu His Ile Pro Leu Gly Thr Pro Asp Asn
15 20 25 30

Phe Cys Thr Tyr

(2) INFORMATION FOR SEQ ID NO: 454:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Placenta

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq VVFMTVAASGASS/FA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 454:

Met Ser Phe Phe Gln Leu Leu Met Lys Arg Lys Glu Leu Ile Pro Leu
-25 -20 -15

Val Val Phe Met Thr Val Ala Ala Ser Gly Ala Ser Ser Phe Ala Val
-10 -5 1

Tyr Ser Leu Trp Lys Thr Asp Val Ile Leu Asp Arg Lys Lys Asn Pro
5 10 15

Glu Pro Trp Glu Thr Val Asp Pro Thr Val Pro Gln Lys Leu Ile Thr
20 25 30 35

Ile Asn Gln Gln Trp Lys Pro Ile Glu Glu Leu Gln Asn Val Gln Arg
40 45 50

Val Thr

(2) INFORMATION FOR SEQ ID NO: 455:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: AMINO ACID

(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq LAHSLLLNEEALA/QI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 455:

```
Met Glu Leu Ala His Ser Leu Leu Leu Asn Glu Glu Ala Leu Ala Gln
-15                 -10                 -5                   1

Ile Thr Glu Ala Lys Arg Pro Val Phe Ile Phe Glu Trp Leu Arg Phe
             5                  10                  15

Leu Asp Lys Val Leu Val Ala Ala Asn Lys Thr Asp Val Lys Glu Lys
        20                  25                  30

Gln Lys Lys Leu Val Glu Gln Leu Thr Gly Leu Ile Ser Ser Ser Pro
    35                  40                  45

Gly Pro Thr Gly
50
```

(2) INFORMATION FOR SEQ ID NO: 456:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 90 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -28..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq LGYLVLSEGAVLA/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 456:

```
Met Thr Ser Ala Leu Thr Gln Gly Leu Glu Arg Ile Pro Asp Gln Leu
            -25                 -20                 -15
```

```
Gly Tyr Leu Val Leu Ser Glu Gly Ala Val Leu Ala Ser Ser Gly Asp
        -10                 -5                  1

Leu Glu Asn Asp Glu Gln Ala Ala Ser Ala Ile Ser Glu Leu Val Ser
  5                 10                  15                  20

Thr Ala Cys Gly Phe Arg Leu His Arg Gly Met Asn Val Pro Phe Lys
                25                  30                  35

Arg Leu Ser Val Val Phe Gly Glu His Thr Leu Leu Val Thr Val Ser
            40                  45                  50

Gly Gln Arg Val Phe Val Val Lys Arg Gly
        55                  60
```

(2) INFORMATION FOR SEQ ID NO: 457:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 148 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -31..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq LVGVLWFVSVTTG/PW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 457:

```
Met Ala Ala Ala Trp Pro Ser Gly Pro Xaa Ala Pro Glu Ala Val Thr
    -30                 -25                 -20

Ala Arg Leu Val Gly Val Leu Trp Phe Val Ser Val Thr Thr Gly Pro
-15                 -10                 -5                   1

Trp Gly Ala Val Ala Thr Ser Ala Gly Gly Glu Glu Ser Leu Lys Cys
            5                   10                  15

Glu Asp Leu Lys Val Gly Gln Tyr Ile Cys Lys Asp Pro Lys Ile Asn
        20                  25                  30

Asp Ala Thr Gln Glu Pro Val Asn Cys Thr Asn Tyr Thr Ala His Val
    35                  40                  45

Ser Cys Phe Pro Ala Pro Asn Ile Thr Cys Lys Asp Xaa Ser Gly Asn
 50                 55                  60                  65

Glu Thr His Phe Thr Gly Asn Glu Val Gly Phe Phe Lys Pro Ile Ser
                70                  75                  80
```

```
Cys Arg Asn Val Asn Gly Tyr Ser Tyr Xaa Xaa Gln Xaa Xaa Val Ser
            85                  90                  95

Phe Ser Trp Met Val Gly Ser Arg Ser Ile Leu Pro Trp Ile Pro Cys
        100                 105                 110

Phe Gly Phe Val
    115
```

(2) INFORMATION FOR SEQ ID NO: 458:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq MVLLTMIARVADG/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 458:

```
Met Val Leu Leu Thr Met Ile Ala Arg Val Ala Asp Gly Leu Pro Leu
            -10                 -5                  1

Ala Ala Ser Met Gln Glu Asp Glu Gln Ser Gly Arg
    5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 459:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.5
seq MVLLTMIARVADG/LP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 459:

```
Met Val Leu Leu Thr Met Ile Ala Arg Val Ala Asp Gly Leu Pro Leu
            -10                 -5                  1

Ala Ala Ser Met Gln Glu Asp Glu Gln Ser Gly Arg Asp Leu Gln Gln
    5                   10                  15

Tyr Gln Ser Gln Ala Lys Gln Leu Phe Arg Lys Leu Asn Glu Gln Ser
20                  25                  30                  35

Pro Thr Arg Cys Thr Leu Glu Ala Gly Ala Met Thr Phe His Tyr Ile
                40                  45                  50

Ile Glu Gln Gly Val Cys Tyr Leu Val Leu Cys Glu Ala Ala Phe Pro
            55                  60                  65

Lys Lys Leu Ala Phe Ala Tyr Leu Glu Asp Leu His Ser Glu Phe Asp
        70                  75                  80

Glu Gln
    85
```

(2) INFORMATION FOR SEQ ID NO: 460:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -69..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.4
seq MMVLSLGIXLASA/SF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 460:

```
Met Thr Ser Gln Pro Val Pro Asn Glu Thr Ile Ile Val Leu Pro Ser
                -65                 -60                 -55

Asn Val Ile Asn Phe Ser Gln Ala Glu Lys Pro Glu Pro Thr Asn Gln
            -50                 -45                 -40

Gly Gln Asp Ser Leu Lys Lys His Leu His Ala Glu Ile Lys Val Ile
        -35                 -30                 -25

Gly Thr Ile Gln Ile Leu Cys Gly Met Met Val Leu Ser Leu Gly Ile
    -20                 -15                 -10
```

Xaa Leu Ala Ser Ala Ser Phe Ser Pro Asn Phe Thr Gln Val Thr Ser
-5 1 5 10

Thr Leu Leu Asn Ser Ala Tyr Pro Phe Ile Gly Pro Phe Phe Phe Ile
15 20 25

Ile Ser Gly Ser Leu Ser Ile
30

(2) INFORMATION FOR SEQ ID NO: 461:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 135 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Placenta

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq AVTSLLSPTPATA/LA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 461:

Met Ala Ser Val Val Leu Ala Leu Arg Thr Arg Thr Ala Val Thr Ser
-25 -20 -15 -10

Leu Leu Ser Pro Thr Pro Ala Thr Ala Leu Ala Val Arg Tyr Ala Ser
-5 1 5

Lys Lys Ser Gly Gly Ser Ser Lys Asn Leu Gly Gly Lys Ser Ser Gly
10 15 20

Arg Arg Gln Gly Ile Lys Lys Met Glu Gly His Tyr Val His Ala Gly
25 30 35

Asn Ile Ile Ala Thr Gln Arg His Phe Arg Trp His Pro Gly Ala His
40 45 50 55

Val Gly Val Gly Lys Xaa Lys Cys Leu Tyr Ala Leu Glu Glu Gly Ile
60 65 70

Val Arg Tyr Thr Lys Glu Val Tyr Val Pro His Pro Arg Asn Thr Glu
75 80 85

Ala Val Xaa Leu Ile Thr Arg Leu Xaa Lys Gly Ala Val Leu Tyr Lys
90 95 100

Thr Phe Val Thr Trp Phe Leu
105 110

(2) INFORMATION FOR SEQ ID NO: 462:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 135 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq AVTSLLSPTPATA/LA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 462:

```
Met Ala Ser Val Val Leu Ala Leu Arg Thr Arg Thr Ala Val Thr Ser
-25                 -20                 -15                 -10

Leu Leu Ser Pro Thr Pro Ala Thr Ala Leu Ala Val Arg Tyr Ala Ser
                -5                  1               5

Lys Lys Ser Gly Gly Ser Ser Lys Asn Leu Gly Gly Lys Ser Ser Gly
        10                  15                  20

Arg Arg Gln Gly Ile Lys Lys Met Glu Gly His Tyr Val His Ala Gly
    25                  30                  35

Asn Ile Ile Ala Thr Gln Arg His Phe Arg Trp His Pro Gly Ala His
40                  45                  50                  55

Val Gly Val Gly Lys Asn Lys Cys Leu Tyr Ala Leu Glu Glu Gly Ile
                60                  65                  70

Xaa Arg Tyr Thr Lys Glu Val Tyr Val Pro His Pro Arg Asn Thr Glu
            75                  80                  85

Ala Val Asp Leu Ile Thr Arg Leu Pro Lys Gly Ala Val Leu Tyr Lys
        90                  95                  100

Thr Phe Val His Val Val Pro
    105                 110
```

(2) INFORMATION FOR SEQ ID NO: 463:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -57..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.3
seq AIALATVLFLIGA/FL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 463:

```
Met Met Pro Ser Arg Thr Asn Leu Ala Thr Gly Ile Pro Ser Ser Lys
        -55                 -50                 -45
Val Lys Tyr Ser Arg Leu Ser Ser Thr Asp Asp Gly Tyr Ile Asp Leu
    -40                 -35                 -30
Gln Phe Lys Lys Thr Pro Pro Lys Ile Pro Tyr Lys Ala Ile Ala Leu
-25                 -20                 -15                 -10
Ala Thr Val Leu Phe Leu Ile Gly Ala Phe Leu Ile Ile Ile Gly Ser
                -5                   1               5
Leu Leu Leu Ser Gly Tyr Ile Ser Lys Gly Gly Ala Asp Arg Ala Val
         10                  15                  20
Pro Val Leu Ile Ile Gly Ile Leu Val Phe Leu Pro Gly Phe Tyr His
     25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 464:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LILSLQVCRPATL/DQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 464:

```
Met Pro Leu Ile Leu Ser Leu Gln Val Cys Arg Pro Ala Thr Leu Asp
```

```
-15                 -10                    -5                    1

Gln Ala Thr Arg Ala Thr Thr Pro Cys Arg Leu Ser Gln Gly Cys Gln
                5                   10                  15

Gln His Pro Thr Xaa Gln
        20
```

(2) INFORMATION FOR SEQ ID NO: 465:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5.6
seq LILSLQVCRPATL/DQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 465:

```
Met Pro Leu Ile Leu Ser Leu Gln Val Cys Arg Pro Ala Thr Leu Asp
-15                 -10                 -5                     1

Gln Ala Thr Arg Ala Thr Thr Pro Cys Arg Leu Ser Gln Gly Cys Gln
                5                   10                  15

Gln His Pro Thr Gln Cys Ser Thr His Leu Gly
        20                  25
```

(2) INFORMATION FOR SEQ ID NO: 466:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -68..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 5.2
seq GVLLLLSSIHFQC/RR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 466:

```
Met Ala Ser Ser Val Gly Asn Val Ala Asp Ser Thr Glu Pro Thr Lys
            -65                 -60                 -55

Arg Met Leu Ser Phe Gln Gly Leu Ala Glu Leu Ala His Arg Glu Tyr
        -50                 -45                 -40

Gln Ala Gly Asp Phe Glu Ala Ala Glu Arg His Cys Met Gln Leu Trp
    -35                 -30                 -25

Arg Gln Glu Pro Asp Asn Thr Gly Val Leu Leu Leu Leu Ser Ser Ile
-20                 -15                 -10                  -5

His Phe Gln Cys Arg Arg Leu Asp Arg Ser Ala His Phe Ser Thr Leu
                  1               5                  10

Ala
```

(2) INFORMATION FOR SEQ ID NO: 467:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 110 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -94..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq VILQLQFLFDVLQ/KT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 467:

```
Met Phe Gly Ser Ala Pro Gln Arg Pro Val Ala Met Thr Thr Ala Gln
                -90                 -85                 -80

Arg Asp Ser Leu Leu Trp Lys Leu Ala Gly Leu Leu Arg Glu Xaa Gly
            -75                 -70                 -65

Asp Val Val Leu Ser Gly Cys Ser Thr Leu Ser Leu Leu Thr Pro Thr
        -60                 -55                 -50

Leu Gln Gln Leu Asn His Val Phe Glu Leu His Leu Gly Pro Trp Gly
    -45                 -40                 -35

Pro Gly Gln Thr Gly Phe Val Ala Leu Pro Ser His Pro Ala Asp Ser
```

```
-30                 -25                 -20                 -15

Pro Val Ile Leu Gln Leu Gln Phe Leu Phe Asp Val Leu Gln Lys Thr
                -10                 -5                  1

Leu Ser Leu Lys Leu Val His Val Ala Gly Pro Gly Pro Thr
        5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 468:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -86..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq LILVGTSKHVAFG/KI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 468:

```
Met Ser Phe Ile Phe Glu Trp Ile Tyr Asn Gly Phe Ser Ser Val Leu
    -85                 -80                 -75

Gln Phe Leu Gly Leu Tyr Lys Lys Ser Gly Lys Leu Val Phe Leu Gly
-70                 -65                 -60                 -55

Leu Asp Asn Ala Gly Lys Thr Thr Leu Leu His Met Leu Lys Asp Asp
                -50                 -45                 -40

Arg Leu Gly Gln His Val Pro Thr Leu His Pro Thr Ser Glu Glu Leu
            -35                 -30                 -25

Thr Ile Ala Gly Met Thr Leu Gln Leu Leu Ile Leu Val Gly Thr Ser
        -20                 -15                 -10

Lys His Val Ala Phe Gly Lys Ile Ile
    -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 469:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -35..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq WYSTVGLLPPVRA/MS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 469:

```
Met Asp Lys Pro Cys Gly Cys Pro Pro Gly Val Cys Asp His Gly Thr
-35                 -30                 -25                 -20

Gly Asp Arg Arg Asp Pro Trp Tyr Ser Thr Val Gly Leu Leu Pro Pro
                -15                 -10                 -5

Val Arg Ala Met Ser Gln Arg Asn Leu Asn
            1               5
```

(2) INFORMATION FOR SEQ ID NO: 470:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Testis

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 5
seq ARALAALVPGVTQ/VD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 470:

```
Met Ala Ala Ala Leu Lys Cys Leu Leu Thr Leu Gly Arg Trp Cys Pro
    -35                 -30                 -25

Gly Leu Gly Val Ala Pro Gln Ala Arg Ala Leu Ala Ala Leu Val Pro
-20                 -15                 -10                 -5

Gly Val Thr Gln Val Asp Asn Lys Ser Gly Phe Leu Gln Lys Arg Pro
                    1               5                   10

His Arg Gln His Pro Gly Ile Leu Lys Leu Pro His Val Arg Leu Pro
        15                  20                  25
```

```
Gln Ala Leu Ala Asn Gly Ala Gln Leu Leu Leu Leu Gly Ser Ala Gly
     30                  35                  40

Pro Thr Met Glu Asn Gln Val Gln Thr Leu Thr Ser Tyr Leu Trp Ser
 45                  50                  55                  60

Arg His Leu Pro Val Glu Pro Xaa Glu Leu Gln Arg Arg Ala Xaa His
                 65                  70                  75

Leu Glu Lys Lys Phe Leu Glu Asn Pro Asp Leu Ser Gln Thr Glu Glu
             80                  85                  90

Lys Leu Arg Gly Ala Gly
         95
```

(2) INFORMATION FOR SEQ ID NO: 471:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -102..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq TVMSALSVAPSKA/RE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 471:

```
Met Val Ala Arg Val Trp Ser Leu Met Arg Phe Leu Ile Lys Gly Ser
        -100                 -95                 -90

Val Ala Gly Gly Ala Val Tyr Leu Val Tyr Asp Gln Glu Leu Leu Gly
    -85                 -80                 -75

Pro Ser Asp Lys Ser Gln Ala Ala Leu Gln Lys Ala Gly Glu Val Val
-70                 -65                 -60                 -55

Pro Pro Ala Met Xaa Gln Phe Ser Gln Tyr Val Cys Gln Gln Thr Gly
                -50                 -45                 -40

Leu Gln Ile Pro Gln Leu Pro Ala Pro Pro Lys Ile Tyr Phe Pro Ile
            -35                 -30                 -25

Arg Asp Ser Trp Xaa Ala Gly Ile Met Thr Val Met Ser Ala Leu Ser
        -20                 -15                 -10

Val Ala Pro Ser Lys Ala Arg Glu Tyr Ser Lys Glu Gly Trp Glu Tyr
    -5                   1                   5                  10
```

Val Lys Ala Leu Gly
15

(2) INFORMATION FOR SEQ ID NO: 472:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq ELQNLXSLQGSQA/CS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 472:

Met Val Asn Glu Leu Gln Asn Leu Xaa Ser Leu Gln Gly Ser Gln Ala
-15 -10 -5

Cys Ser Ser Ser Lys Gln Arg Phe
1 5

(2) INFORMATION FOR SEQ ID NO: 473:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -24..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq FFFSIQPFLPCSS/RP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 473:

Met Leu Tyr Met Ser Leu Lys Tyr Ile Arg Ala Phe Phe Phe Ser Ile

```
                -20                  -15                  -10

Gln Pro Phe Leu Pro Cys Ser Ser Arg Pro Leu Lys Ser Pro Ser Pro
            -5                   1               5

Val Ala His Pro Thr Asn Ile Ser Val Ser Glu Asn Ala Gln Arg Cys
    10                  15                  20

Leu Xaa Thr Ser Pro Trp
 25                 30
```

(2) INFORMATION FOR SEQ ID NO: 474:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -79..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq WVIVLTSWITIFQ/IY

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 474:

```
Met Asn Leu Glu Arg Val Ser Asn Glu Glu Lys Leu Asn Leu Cys Arg
                -75                 -70                 -65

Lys Tyr Tyr Leu Gly Gly Phe Ala Phe Leu Pro Phe Leu Trp Leu Val
            -60                 -55                 -50

Asn Ile Phe Trp Phe Phe Arg Glu Ala Phe Leu Val Pro Ala Tyr Thr
        -45                 -40                 -35

Glu Gln Ser Gln Ile Lys Gly Tyr Val Trp Arg Ser Ala Val Gly Phe
    -30                 -25                 -20

Leu Phe Trp Val Ile Val Leu Thr Ser Trp Ile Thr Ile Phe Gln Ile
-15                 -10                  -5                   1

Tyr Arg Pro Arg Trp Gly Ala Leu Gly Asp Xaa Leu Ser Phe Thr Ile
             5                  10                  15

Pro Leu Gly Thr Pro Asp Asn Phe Cys Thr Tyr
        20                  25
```

(2) INFORMATION FOR SEQ ID NO: 475:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -70..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq LVFVLLFIFVKRQ/IM

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 475:

```
Met Ala Gly Glu Leu Gln Gly Thr Gln Ala Pro Ser Leu Arg Gly Xaa
-70                 -65                 -60                 -55

Gly Leu Thr Ser Gln Asp Ser Gly Val Asn Pro Asn Asn Ser Xaa Arg
                -50                 -45                 -40

Gly Arg Glu Ala Met Ala Ser Gly Ser Asn Trp Leu Ser Gly Val Asn
            -35                 -30                 -25

Val Val Leu Val Met Ala Tyr Gly Ser Leu Val Phe Val Leu Leu Phe
        -20                 -15                 -10

Ile Phe Val Lys Arg Gln Ile Met Arg Phe Ala Met Lys Ser Arg Arg
    -5                   1               5                  10

Gly Pro His Val Pro Val Gly Xaa Gln Cys Pro Gln Xaa Cys Tyr Asn
                15                  20                  25

Tyr Leu Tyr
```

(2) INFORMATION FOR SEQ ID NO: 476:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -56..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9

seq FACVPGASPTTLA/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 476:

```
Met Thr Gly Phe Leu Leu Pro Pro Ala Ser Arg Gly Thr Arg Arg Ser
    -55                 -50                 -45

Cys Ser Arg Ser Arg Lys Arg Gln Thr Arg Arg Arg Arg Asn Pro Ser
-40                 -35                 -30                 -25

Ser Phe Val Ala Ser Cys Pro Thr Leu Leu Pro Phe Ala Cys Val Pro
                -20                 -15                 -10

Gly Ala Ser Pro Thr Thr Leu Ala Phe Pro Pro Val Val Leu Thr Gly
            -5                  1               5

Pro Ser Thr Asp Gly Ile Pro Phe Ala Leu Ser Leu Gln Arg Val Pro
    10                  15                  20

Phe Val
25
```

(2) INFORMATION FOR SEQ ID NO: 477:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.9
seq VLCTNQVLITARA/VP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 477:

```
Met Glu Glu Xaa Ser Xaa Pro Leu Val Glu Phe Val Lys Val Leu Cys
    -25                 -20                 -15

Thr Asn Gln Val Leu Ile Thr Ala Arg Ala Val Pro Thr Lys Lys Ala
-10                 -5                  1                   5

Ser Val Arg Cys Val Xaa Lys Arg Phe Trp Ile Pro Lys Thr Thr Ser
            10                  15                  20

Lys His Leu Ser Arg Cys Ile Asp Gly Ile Ser Gly Phe Leu Asn Asp
        25                  30                  35
```

```
Phe Thr Phe Cys Leu Glu Phe Ser Arg His Arg Cys
     40                 45                  50
```

(2) INFORMATION FOR SEQ ID NO: 478:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq LXXVVAFVAPGES/QQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 478:

```
Met Val Arg Arg Leu Xaa Xaa Val Val Ala Phe Val Ala Pro Gly Glu
        -15                 -10                 -5

Ser Gln Gln Glu Glu Pro Pro Thr Asp Asn Gln Asp Ile Glu Pro Gly
  1                   5                  10                  15

Gln Glu Arg Glu Gly Thr Pro Pro Ile Glu Glu Arg Lys Val Glu Gly
                 20                  25                  30

Asp Cys Gln Glu Met Asp Leu Glu Lys Thr Arg Ser Glu Arg Gly Asp
             35                  40                  45

Gly Ser Asp Val Lys Glu Lys Thr Pro Pro Asn Xaa Lys His Ala Lys
         50                  55                  60

Thr Lys Glu Ala Gly Asp Gly Pro Leu
     65                  70
```

(2) INFORMATION FOR SEQ ID NO: 479:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 149 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -37..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq PIVRLLSCPGTVA/KD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 479:

```
Met Ala Val Pro Gly Val Gly Leu Leu Thr Arg Leu Asn Leu Cys Ala
        -35                 -30                 -25

Arg Arg Arg Thr Arg Val Gln Arg Pro Ile Val Arg Leu Leu Ser Cys
    -20                 -15                 -10

Pro Gly Thr Val Ala Lys Asp Leu Arg Arg Asp Glu Gln Pro Ser Gly
 -5                   1               5                  10

Ser Val Glu Thr Gly Phe Glu Asp Lys Ile Pro Lys Arg Arg Phe Ser
             15                  20                  25

Glu Met Gln Asn Glu Arg Arg Glu Gln Ala Gln Arg Thr Val Leu Ile
         30                  35                  40

His Cys Pro Glu Lys Ile Ser Glu Asn Lys Phe Xaa Lys Tyr Leu Ser
     45                  50                  55

Gln Phe Gly Pro Ile Asn Asn His Phe Phe Tyr Glu Ser Phe Gly Leu
 60                  65                  70                  75

Tyr Ala Val Val Glu Phe Cys Gln Lys Glu Ser Ile Gly Ser Leu Gln
                 80                  85                  90

Asn Gly Thr His Thr Pro Ser Thr Ala Met Glu Thr Ala Ile Pro Phe
             95                 100                 105

Arg Ser Arg Ser Ser
        110
```

(2) INFORMATION FOR SEQ ID NO: 480:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -60..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq LVILSLKSQTLDA/ET

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 480:

```
Met Met Ala Ala Val Pro Pro Gly Leu Glu Pro Trp Asn Arg Val Arg
-60                 -55                 -50                 -45

Ile Pro Lys Ala Gly Asn Arg Ser Ala Val Thr Val Gln Asn Pro Gly
                -40                 -35                 -30

Ala Ala Leu Asp Leu Cys Ile Ala Ala Val Ile Lys Glu Cys His Leu
            -25                 -20                 -15

Val Ile Leu Ser Leu Lys Ser Gln Thr Leu Asp Ala Glu Thr Asp Val
        -10                 -5                  1

Leu Cys Ala Val Leu Tyr Ser Asn His Asn Arg Met Gly Arg His Lys
5                   10                  15                  20

Pro His Leu Ala Leu Lys Gln Val Glu Gln Cys Leu Lys Arg Leu Xaa
                25                  30                  35

Asn Met Asn Leu Glu Gly Gly
            40
```

(2) INFORMATION FOR SEQ ID NO: 481:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -33..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq SLVHLLCQNQVLG/NP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 481:

```
Met Ala Ser Leu Asp Arg Val Lys Val Leu Val Leu Gly Asp Ser Gly
            -30                 -25                 -20

Val Gly Lys Ser Ser Leu Val His Leu Leu Cys Gln Asn Gln Val Leu
        -15                 -10                 -5

Gly Asn Pro Ser Trp Thr Val Gly Cys Ser Val Asp Val Arg Val His
    1               5                   10                  15

Asp Tyr Lys Glu Gly Thr Pro Glu Glu Lys Thr Tyr Tyr Ile Glu Leu
                20                  25                  30
```

```
Trp Asp Val Gly Gly Ser Val Gly Ser Ala Ser Ser Val Lys Ser Thr
            35                  40                  45

Arg Ala Val Phe Tyr Asn Ser Val Asn Gly Ile Ile Xaa Val His Asp
        50                  55                  60

Leu Thr Xaa Gly Lys Ser Ser Gln Xaa Leu Arg Arg Trp Ser Leu Glu
    65                  70                  75

Ala Leu Asn Arg Asp Leu Val Pro Thr Gly Val Leu Val
80                  85                  90
```

(2) INFORMATION FOR SEQ ID NO: 482:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -31..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.8
seq WAFSCGTWLPSRA/EW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 482:

```
Met Val Phe Pro Ala Lys Arg Phe Cys Leu Val Pro Ser Met Glu Gly
    -30                 -25                 -20

Val Arg Trp Ala Phe Ser Cys Gly Thr Trp Leu Pro Ser Arg Ala Glu
-15                 -10                 -5                   1

Trp Leu Leu Xaa Val Arg Ser Ile Gln Pro Glu Glu Lys Glu Arg Ile
            5                   10                  15

Gly Gln Phe Val Phe Ala Arg Asp Ala Lys Ala Ala Met Ala Gly Arg
        20                  25                  30

Leu Met Ile Arg Lys Leu Val Ala Glu Asn Arg
    35                  40
```

(2) INFORMATION FOR SEQ ID NO: 483:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -26..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LIMQLGSVLLTRC/PF

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 483:

```
Met Ala Ser Lys Ile Gly Ser Arg Arg Trp Met Leu Gln Leu Ile Met
    -25                 -20                 -15

Gln Leu Gly Ser Val Leu Leu Thr Arg Cys Pro Phe Trp Gly Cys Phe
-10                 -5                  1                   5

Ser Gln Leu Met Leu Tyr Ala Glu Arg Ala Glu Ala Arg Arg Lys Pro
            10                  15                  20

Asp Ile Pro Val Pro Tyr Leu Tyr Phe Asp Met Gly Ala Ala Val Leu
        25                  30                  35

Cys Ala Arg
    40
```

(2) INFORMATION FOR SEQ ID NO: 484:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -31..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LAVDSWWLDPGHA/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 484:

```
Met Leu Ser Lys Gly Leu Lys Arg Lys Arg Glu Glu Glu Glu Glu Lys
    -30                 -25                 -20

Glu Pro Leu Ala Val Asp Ser Trp Trp Leu Asp Pro Gly His Ala Ala
```

-15 -10 -5 1

Val Ala Gln Ala Pro Pro Ala Val Ala Ser Ser Ser Leu Phe Asp Leu
5 10 15

Ser Val Leu Lys Leu His His Ser Arg Gly
20 25

(2) INFORMATION FOR SEQ ID NO: 485:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq SLAAALTLHGHWG/LG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 485:

Met Asp Tyr Ser Leu Ala Ala Ala Leu Thr Leu His Gly His Trp Gly
-15 -10 -5

Leu Gly Gln Val Val Thr Asp Tyr Val His Gly Asp Ala Leu Gln Lys
1 5 10 15

Ala

(2) INFORMATION FOR SEQ ID NO: 486:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -72..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LSLXASYIFGISG/FE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 486:

```
Met Ser Tyr Ile Thr Ser Gln Glu Met Lys Cys Ile Leu His Trp Phe
        -70                 -65                 -60

Ala Asn Trp Ser Gly Pro Gln Arg Glu Arg Phe Leu Glu Asp Leu Val
    -55                 -50                 -45

Ala Lys Ala Val Pro Glu Lys Leu Gln Pro Xaa Leu Asp Ser Leu Glu
-40                 -35                 -30                 -25

Gln Leu Ser Val Ser Gly Ala Asp Asp His Leu Leu Ser Leu Xaa Ala
                -20                 -15                 -10

Ser Tyr Ile Phe Gly Ile Ser Gly Phe Glu Ala Gly Ala Glu Gln Glu
            -5                   1               5

Arg Asn Glu Phe Val Arg Gln Ser
    10                  15
```

(2) INFORMATION FOR SEQ ID NO: 487:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -76..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq LIVYLWVVSFIAS/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 487:

```
Met Pro Leu Leu Cys Gln Ile Glu Met Glu Tyr Leu Leu Leu Lys Trp
    -75                 -70                 -65

Gln Met Thr Met Leu Gln Ser Met Leu Cys Asp Leu Val Ser Tyr Pro
-60                 -55                 -50                 -45

Leu Leu Pro Leu Gln Gln Thr Lys Glu Ala Asn Leu Asp Phe Pro Lys
                -40                 -35                 -30

Ile Lys Val Ser Ser Val Thr Ile Thr Pro Thr Arg Trp Phe Xaa Leu
            -25                 -20                 -15
```

```
Ile Val Tyr Leu Trp Val Val Ser Phe Ile Ala Ser Ser Ser
    -10                     -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 488:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 67 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Surrenals

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.7
seq SVMGVCLLIPGLA/TA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 488:

```
Met Trp Phe Glu Ile Leu Pro Gly Leu Ser Val Met Gly Val Cys Leu
    -20                     -15                 -10

Leu Ile Pro Gly Leu Ala Thr Ala Tyr Ile His Xaa Phe Thr Asn Arg
    -5                  1               5                   10

Gly Lys Glu Lys Arg Val Ala His Phe Gly Tyr His Trp Ser Leu Met
                15                  20                  25

Glu Arg Asp Arg Arg Ile Ser Gly Val Asp Arg Tyr Tyr Val Ser Lys
            30                  35                  40

Gly Pro Gly
        45
```

(2) INFORMATION FOR SEQ ID NO: 489:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide

(B) LOCATION: -46..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq LLVSLVLRXPAKS/TR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 489:

```
Met Glu Phe Lys Leu Glu Ala His Arg Ile Val Ser Ile Ser Leu Gly
    -45                 -40                 -35

Lys Ile Tyr Asn Ser Arg Val Gln Arg Gly Gly Ile Lys Leu His Lys
-30                 -25                 -20                 -15

Asn Leu Leu Val Ser Leu Val Leu Arg Xaa Pro Ala Lys Ser Thr Arg
                -10                 -5                   1

Ala Gly
```

(2) INFORMATION FOR SEQ ID NO: 490:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 109 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -97..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq IASGLGLXLDCWT/SS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 490:

```
Met Ala Val Leu Ser Lys Glu Tyr Gly Phe Val Leu Leu Thr Gly Ala
        -95                 -90                 -85

Ala Ser Phe Ile Met Val Ala His Leu Ala Ile Asn Val Ser Lys Ala
    -80                 -75                 -70

Arg Lys Lys Tyr Lys Val Glu Tyr Pro Ile Met Tyr Ser Thr Asp Pro
-65                 -60                 -55                 -50

Glu Asn Gly His Ile Phe Asn Cys Ile Gln Arg Ala His Gln Asn Thr
                -45                 -40                 -35

Leu Glu Val Tyr Pro Xaa Phe Leu Phe Phe Leu Ala Val Gly Gly Val
            -30                 -25                 -20

Tyr His Pro Arg Ile Ala Ser Gly Leu Gly Leu Xaa Leu Asp Cys Trp
        -15                 -10                 -5
```

Thr Ser Ser Leu Cys Leu Trp Leu Leu His Gly Pro Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO: 491:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -42..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.6
seq RIPSLPGSPVCWA/WP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 491:

Met Asp Gly His Trp Ser Ala Ala Phe Ser Ala Leu Thr Val Thr Ala
-40 -35 -30

Met Ser Ser Trp Ala Arg Arg Arg Ser Ser Ser Ser Arg Arg Ile Pro
-25 -20 -15

Ser Leu Pro Gly Ser Pro Val Cys Trp Ala Trp Pro Trp
-10 -5 1

(2) INFORMATION FOR SEQ ID NO: 492:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Liver

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -16..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq RLLLRRFLASVIS/RK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 492:

```
Met Ala Gln Arg Leu Leu Leu Arg Arg Phe Leu Ala Ser Val Ile Ser
    -15                 -10                 -5

Arg Lys Pro Ser Gln Gly Gln Trp Pro Pro Leu Thr Ser Arg Ala Leu
 1               5                  10                  15

Gln Thr Pro Gln Cys Ser Pro Gly Gly Leu Thr Val Thr Pro Asn Pro
            20                  25                  30

Ala Arg Thr
        35
```

(2) INFORMATION FOR SEQ ID NO: 493:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq FLLLLEVSHLLLI/IN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 493:

```
Met Ala Ser Leu Lys Pro Ala Phe Val Asn Tyr Phe Phe Leu Leu Leu
-25                 -20                 -15                 -10

Leu Glu Val Ser His Leu Leu Leu Ile Ile Asn Ala Glu Gly
                -5                  1               5
```

(2) INFORMATION FOR SEQ ID NO: 494:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: -77..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.5
seq LFWVIVLTSWITI/FQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 494:

```
Met Asn Leu Glu Arg Val Ser Asn Glu Glu Lys Leu Asn Leu Cys Arg
    -75                 -70                 -65

Lys Tyr Tyr Leu Gly Gly Phe Ala Phe Leu Pro Phe Leu Trp Leu Val
    -60                 -55                 -50

Asn Ile Phe Trp Phe Phe Arg Glu Ala Phe Leu Val Pro Ala Tyr Thr
-45                 -40                 -35                 -30

Glu Gln Ser Gln Ile Lys Gly Tyr Val Trp Arg Ser Ala Val Gly Phe
                -25                 -20                 -15

Leu Phe Trp Val Ile Val Leu Thr Ser Trp Ile Thr Ile Phe Gln Ile
            -10                 -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 495:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lung (cells)

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -21..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq AVASSFFCASLFS/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 495:

```
Met Ala Gln Leu Gly Ala Val Val Ala Val Ala Ser Ser Phe Phe Cys
    -20                 -15                 -10

Ala Ser Leu Phe Ser Ala Val His Lys Ile Glu Glu Gly His Ile Gly
-5                  1               5                   10

Val Tyr Tyr Arg Gly Gly Val
            15
```

(2) INFORMATION FOR SEQ ID NO: 496:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 64 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq LVFMVPLVGLIHL/GW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 496:

```
Met Ser Leu Arg Asn Leu Trp Arg Asp Tyr Lys Val Leu Val Phe Met
-25                 -20                 -15                 -10

Val Pro Leu Val Gly Leu Ile His Leu Gly Trp Tyr Arg Ile Lys Ser
                 -5                   1               5

Ser Pro Val Phe Gln Ile Pro Lys Asn Asp Asp Ile Pro Glu Gln Asp
        10                  15                  20

Ser Leu Gly Leu Ser Asn Leu Gln Lys Ser Gln Ile Gln Gly Ile Leu
    25                  30                  35
```

(2) INFORMATION FOR SEQ ID NO: 497:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Spleen

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -23..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq VFCLLISIPTPSA/HL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 497:

```
Met Gly Trp Asp Gly Cys Lys Cys Leu Gly Val Phe Cys Leu Leu Ile
```

```
        -20                  -15                 -10

Ser Ile Pro Thr Pro Ser Ala His Leu
        -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 498:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 149 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -118..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq ILAHRLGLIPIHA/DP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 498:

```
Met Ala Ala Ser Gln Ala Val Glu Glu Met Arg Thr Ala Trp Phe Trp
            -115                -110                -105

Gly Ser Leu Gly Phe Ala Met Ser Ile Leu Leu Thr Phe Pro Val Thr
        -100                -95                 -90

Ile Pro Val Met Met Met Pro Gly Thr Arg Xaa Gly Phe Glu Xaa Arg
    -85                 -80                 -75

Xaa Phe Arg Val Asp Val Val His Met Asp Glu Asn Ser Leu Glu Phe
-70                 -65                 -60                 -55

Asp Met Val Gly Ile Asp Ala Ala Ile Ala Asn Ala Phe Arg Arg Ile
                -50                 -45                 -40

Leu Leu Ala Glu Val Pro Thr Met Ala Val Glu Lys Val Leu Val Tyr
            -35                 -30                 -25

Asn Asn Thr Ser Ile Val Gln Asp Glu Ile Leu Ala His Arg Leu Gly
        -20                 -15                 -10

Leu Ile Pro Ile His Ala Asp Pro Arg Leu Phe Glu Tyr Arg Asn Gln
    -5                  1               5                   10

Gly Asp Glu Glu Gly Thr Glu Ile Asp Thr Leu Gln Phe Arg Leu Gln
                15                  20                  25

Val Arg Cys Thr Arg
            30
```

(2) INFORMATION FOR SEQ ID NO: 499:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -77..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq FEARIALLPLLQA/ET

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 499:

```
Met Ala Ala Ser Lys Val Lys Gln Asp Met Pro Pro Pro Gly Gly Tyr
        -75                 -70                 -65

Gly Pro Ile Asp Tyr Lys Arg Asn Leu Pro Arg Arg Gly Leu Ser Gly
    -60                 -55                 -50

Tyr Ser Met Leu Ala Ile Gly Ile Gly Thr Leu Ile Tyr Gly His Trp
-45                 -40                 -35                 -30

Ser Ile Met Lys Trp Asn Arg Glu Arg Arg Arg Leu Gln Ile Glu Asp
                -25                 -20                 -15

Phe Glu Ala Arg Ile Ala Leu Leu Pro Leu Leu Gln Ala Glu Thr Asp
            -10                 -5                   1

Arg Xaa Thr Leu Gln Met Leu Arg Glu Asn Leu Glu Glu Glu Ala Ile
    5                   10                  15

Ile Met Xaa Asp Val Xaa Asp Trp Xaa Val Gly Xaa Xaa Xaa Val Pro
20                  25                  30                  35

His Asn Pro Leu Gly Ala Pro Leu Asp Arg Gly Ala
                40                  45
```

(2) INFORMATION FOR SEQ ID NO: 500:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 49 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -42..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq VLFFTGWWIIIDA/AV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 500:

```
Met Ser Gly Phe Leu Glu Gly Leu Arg Cys Ser Glu Cys Ile Asp Trp
        -40                 -35                 -30

Gly Glu Lys Arg Asn Thr Ile Ala Ser Ile Ala Ala Gly Val Leu Phe
    -25                 -20                 -15

Phe Thr Gly Trp Trp Ile Ile Ile Asp Ala Ala Val Ile Tyr Pro Thr
-10                 -5                  1               5

Arg
```

(2) INFORMATION FOR SEQ ID NO: 501:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 58 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -44..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq LVFLTFLSIPSFV/GL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 501:

```
Met Met Thr Gln Glu Pro Gly Ile Tyr Thr Trp Pro Glu Lys Thr Arg
                -40                 -35                 -30

Ile Ile Cys Ser Ala Cys Ser Ser Val Pro Leu Pro Trp Thr Val Leu
            -25                 -20                 -15

Val Phe Leu Thr Phe Leu Ser Ile Pro Ser Phe Val Gly Leu Arg Asn
        -10                 -5                  1

Ile Arg Ala Glu Thr Phe Leu Gln Asn Val
  5                 10
```

(2) INFORMATION FOR SEQ ID NO: 502:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -14..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.4
seq FLTALLWRGRIPG/RQ

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 502:

```
Met Phe Leu Thr Ala Leu Leu Trp Arg Gly Arg Ile Pro Gly Arg Gln
                -10                 -5                   1

Trp Ile Gly Lys His Arg Arg Pro Arg Phe Val Ser Leu Arg Ala Lys
         5                   10                  15

Gln Asn Met Ile Arg Arg Leu Glu Ile Glu Ala Glu Asn His Tyr Trp
     20                  25                  30

Leu Ser Met Pro Tyr Met Thr Arg Glu Gln Glu Arg Gly His Ala Xaa
 35                  40                  45                  50

Leu Arg Arg Arg Glu Ala Phe Glu Ala Ile Lys Ala Ala Ala Thr Ser
                 55                  60                  65

Lys Phe Pro Pro His Arg Phe Ile Ala Asp Gln Leu Asp His Leu Xaa
             70                  75                  80

Xaa His Gln Glu Met Val Leu Ile Leu Ser Arg His Pro Trp Ile Leu
         85                  90                  95

Trp Ile Thr Glu Leu Thr Ile Phe Thr Trp Ser Gly Leu Lys Asn Cys
    100                 105                 110

Ser Leu Cys Glu Asn Glu Leu Trp Thr Ser Leu Tyr
115                 120                 125
```

(2) INFORMATION FOR SEQ ID NO: 503:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -90..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq TCLTACWTALCCC/CL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 503:

```
Met Asn Gln Glu Asn Pro Pro Pro Tyr Pro Gly Pro Gly Pro Thr Ala
-90                 -85                 -80                 -75

Pro Tyr Pro Pro Tyr Pro Pro Gln Pro Met Gly Pro Gly Xaa Met Gly
                -70                 -65                 -60

Gly Pro Tyr Pro Pro Pro Gln Gly Tyr Pro Tyr Gln Gly Tyr Pro Gln
            -55                 -50                 -45

Tyr Gly Trp Gln Gly Gly Pro Gln Glu Pro Pro Lys Thr Thr Val Tyr
        -40                 -35                 -30

Val Val Glu Asp Gln Arg Arg Asp Glu Leu Gly Pro Ser Thr Cys Leu
    -25                 -20                 -15

Thr Ala Cys Trp Thr Ala Leu Cys Cys Cys Cys Leu Trp
-10                 -5                   1
```

(2) INFORMATION FOR SEQ ID NO: 504:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -54..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.3
seq LIVWLLVKSFSES/GI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 504:

```
Met Ala Ser Leu Glu Val Ser Arg Ser Pro Arg Arg Ser Arg Arg Glu
                -50                 -45                 -40

Leu Glu Val Arg Ser Pro Arg Gln Asn Lys His Ser Val Leu Leu Pro
            -35                 -30                 -25

Thr Tyr Asn Glu Arg Glu Glu Leu Pro Leu Ile Val Trp Leu Leu Val
        -20                 -15                 -10

Lys Ser Phe Ser Glu Ser Gly Ile Asn Tyr Glu Ile Ile Ile Ile Asp
    -5                  1               5                   10

Asp Gly Ser Pro Asp Gly Thr Arg Asp Val Ala Glu Gln Leu Glu Lys
                15                  20                  25

Ile Tyr Gly Ser Asp Arg Ile Leu Leu Arg Pro Arg Glu Lys Lys Leu
            30                  35                  40

Gly Leu Gly Thr Ala Tyr Ile His Gly Met Xaa Thr Cys His Arg Xaa
        45                  50                  55

Leu His His Tyr Tyr Gly Cys
    60                  65
```

(2) INFORMATION FOR SEQ ID NO: 505:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 amino acids
- (B) TYPE: AMINO ACID
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
- (A) ORGANISM: Homo Sapiens
- (F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
- (A) NAME/KEY: sig_peptide
- (B) LOCATION: -14..-1
- (C) IDENTIFICATION METHOD: Von Heijne matrix
- (D) OTHER INFORMATION: score 4.2
  seq CPTCLCAPSXXWG/EP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 505:

```
Met Cys Pro Thr Cys Leu Cys Ala Pro Ser Xaa Xaa Trp Gly Glu Pro
                -10                 -5                      1

Val Gly Ser Pro Gly Leu Ser Ser Pro Val Leu Ser Pro Ser Lys Lys
        5                   10                  15

Ala Arg Ser
    20
```

(2) INFORMATION FOR SEQ ID NO: 506:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq AVAASAASGQAEG/KK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 506:

```
Met Ala Ala Ala Thr Gly Ala Val Ala Ala Ser Ala Ala Ser Gly Gln
            -15                 -10                 -5

Ala Glu Gly Lys Lys Ile Thr Asp Leu Arg Val Ile Asp Leu Lys Ser
             1               5                  10

Glu Leu Lys Arg Arg Asn Leu Asp Ile Thr Gly Val Lys Thr Val Leu
    15                  20                  25

Ile Ser Arg Leu Arg
30
```

(2) INFORMATION FOR SEQ ID NO: 507:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 137 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -17..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq SLLXRVSVTAVAA/LS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 507:

```
Met Ala Ala Met Ser Leu Leu Xaa Arg Val Ser Val Thr Ala Val Ala
```

-15 -10 -5

Ala Leu Ser Gly Arg Pro Leu Gly Thr Xaa Leu Gly Phe Gly Gly Phe
1 5 10 15

Leu Thr Arg Gly Phe Pro Lys Ala Ala Ala Pro Val Arg His Ser Gly
20 25 30

Asp His Gly Lys Arg Leu Phe Val Ile Arg Pro Ser Arg Phe Tyr Asp
35 40 45

Arg Arg Phe Leu Lys Leu Leu Arg Phe Tyr Ile Ala Leu Thr Gly Ile
50 55 60

Pro Val Ala Xaa Phe Ile Thr Leu Val Asn Val Phe Ile Gly Gln Ala
65 70 75

Glu Leu Ala Glu Ile Pro Glu Gly Tyr Val Pro Glu His Trp Glu Tyr
80 85 90 95

Tyr Lys His Pro Ile Ser Arg Trp Ile Ala Arg Asn Phe Tyr Asp Ser
100 105 110

Pro Xaa Lys Ile Tyr Glu Arg Thr Met
115 120

(2) INFORMATION FOR SEQ ID NO: 508:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 131 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.2
seq LDLLRGLPRVSLA/NL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 508:

Met Ala Gly Pro Leu Gln Gly Gly Gly Ala Arg Ala Leu Asp Leu Leu
-25 -20 -15 -10

Arg Gly Leu Pro Arg Val Ser Leu Ala Asn Leu Lys Pro Asn Pro Gly
-5 1 5

Ser Lys Lys Pro Glu Arg Arg Pro Arg Gly Arg Arg Arg Gly Arg Lys
10 15 20

Cys Gly Arg Gly His Lys Gly Glu Arg Gln Arg Gly Thr Arg Pro Arg

```
25                  30                  35

Leu Gly Phe Glu Gly Gly Gln Thr Pro Phe Tyr Ile Arg Xaa Pro Lys
 40                  45                  50                  55

Tyr Gly Phe Asn Glu Gly His Ser Phe Arg Arg Gln Tyr Lys Pro Leu
                60                  65                  70

Ser Leu Asn Arg Leu Gln Tyr Leu Ile Asp Leu Gly Arg Val Asp Pro
            75                  80                  85

Ser Gln Pro Ile Asp Leu Thr Gln Leu Val Asn Gly Arg Gly Val Thr
        90                  95                 100

Ile Ala Pro
    105
```

(2) INFORMATION FOR SEQ ID NO: 509:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 136 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -41..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq GILILWIIRLLFS/KT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 509:

```
Met Ala Thr Ala Thr Glu Gln Trp Val Leu Val Glu Met Val Gln Ala
    -40                 -35                 -30

Leu Tyr Glu Ala Pro Ala Tyr His Leu Ile Leu Glu Gly Ile Leu Ile
-25                 -20                 -15                 -10

Leu Trp Ile Ile Arg Leu Leu Phe Ser Lys Thr Tyr Lys Leu Gln Glu
                -5                   1                   5

Arg Ser Asp Leu Thr Val Lys Glu Lys Glu Glu Leu Ile Glu Glu Trp
        10                  15                  20

Gln Pro Glu Pro Leu Val Pro Pro Val Pro Lys Asp His Pro Ala Leu
    25                  30                  35

Asn Tyr Asn Ile Val Ser Gly Pro Pro Ser His Lys Thr Val Val Asn
 40                  45                  50                  55
```

Gly Lys Glu Cys Ile Asn Phe Ala Ser Phe Asn Phe Leu Gly Leu Leu
60 65 70

Asp Asn Pro Arg Val Lys Ala Ala Ala Leu Ala Ser Leu Lys Lys Tyr
75 80 85

Gly Val Gly Thr Cys Gly Pro Cys
90 95

(2) INFORMATION FOR SEQ ID NO: 510:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -79..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq QGVLFICFTCARS/FP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 510:

Met Glu Asp Pro Asn Pro Glu Glu Asn Met Xaa Gln Gln Asp Ser Pro
-75 -70 -65

Lys Glu Arg Ser Pro Gln Ser Pro Gly Gly Asn Ile Cys His Leu Gly
-60 -55 -50

Ala Pro Lys Cys Thr Arg Cys Leu Ile Thr Phe Ala Asp Ser Lys Xaa
-45 -40 -35

Xaa Glu Arg His Met Lys Arg Glu His Pro Ala Asp Phe Val Ala Gln
-30 -25 -20

Lys Leu Gln Gly Val Leu Phe Ile Cys Phe Thr Cys Ala Arg Ser Phe
-15 -10 -5 1

Pro Ser

(2) INFORMATION FOR SEQ ID NO: 511:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 142 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -32..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq RLLSSLLLTMSNN/NP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 511:

```
Met Asn Val Ile Asp His Val Arg Asp Met Ala Ala Ala Gly Leu His
        -30                 -25                 -20

Ser Asn Val Arg Leu Leu Ser Ser Leu Leu Leu Thr Met Ser Asn Asn
    -15                 -10                 -5

Asn Pro Glu Leu Phe Ser Pro Pro Gln Lys Tyr Gln Leu Leu Val Tyr
  1               5                  10                  15

His Ala Asp Ser Leu Phe His Asp Lys Glu Tyr Arg Asn Ala Val Ser
            20                  25                  30

Lys Tyr Thr Met Ala Leu Gln Gln Lys Lys Ala Leu Ser Lys Thr Ser
        35                  40                  45

Lys Val Arg Pro Ser Thr Gly Asn Ser Ala Ser Thr Pro Gln Ser Gln
    50                  55                  60

Cys Leu Pro Ser Glu Ile Glu Val Lys Tyr Lys Met Ala Glu Cys Tyr
 65                  70                  75                  80

Thr Met Leu Lys Gln Asp Lys Asp Ala Ile Ala Ile Leu Asp Gly Xaa
                85                  90                  95

Pro Phe Lys Thr Lys Asn Ser Gln Asn Lys His Asp Ala Gly
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 512:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -58..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix

(D) OTHER INFORMATION: score 4.1
seq LVHHCPTWQWATG/EE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 512:

```
Met Gln Asn Val Ile Asn Thr Val Lys Gly Lys Ala Leu Glu Val Ala
            -55                 -50                 -45

Glu Tyr Leu Thr Pro Val Leu Lys Glu Ser Lys Phe Lys Glu Thr Gly
        -40                 -35                 -30

Val Ile Thr Pro Glu Glu Phe Val Ala Ala Gly Asp His Leu Val His
    -25                 -20                 -15

His Cys Pro Thr Trp Gln Trp Ala Thr Gly Glu Glu Leu Lys Val Lys
-10                 -5                  1                   5

Ala Tyr Leu Pro Thr Gly Lys Trp
            10
```

(2) INFORMATION FOR SEQ ID NO: 513:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 91 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Colon

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -88..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4.1
seq CIQRLPWLLLCRG/IT

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 513:

```
Met Ala Thr Leu Thr Phe Ser Leu Arg Lys Pro Leu Gln Arg Ser Leu
            -85                 -80                 -75

Ile Arg Pro Ser His Leu Pro Leu Cys Cys Phe Asp Trp Arg Leu Ser
        -70                 -65                 -60

His Tyr Tyr Arg Leu Pro Pro Ala Val Arg Leu His Gln Gln Arg Gly
    -55                 -50                 -45

Gly Arg Pro Gly Arg Ser Ser Ala Asp His Trp His Ser Gly Val Pro
-40                 -35                 -30                 -25

Thr Arg Ile Leu Pro Pro Ala His Arg Leu Leu Cys Ile Gln Arg Leu
                -20                 -15                 -10
```

Pro Trp Leu Leu Leu Cys Arg Gly Ile Thr Ser
-5 1

(2) INFORMATION FOR SEQ ID NO: 514:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -49..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq PSLAAGLLFGSXA/GL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 514:

Met Glu Lys Pro Leu Phe Pro Leu Val Pro Leu His Trp Phe Gly Phe
-45 -40 -35

Gly Tyr Thr Ala Leu Val Val Ser Gly Gly Ile Val Gly Tyr Val Lys
-30 -25 -20

Thr Gly Ser Val Pro Ser Leu Ala Ala Gly Leu Leu Phe Gly Ser Xaa
-15 -10 -5

Ala Gly Leu Gly Ala Tyr Gln Leu Tyr Gln Asp Pro Arg Asn Val Trp
1 5 10 15

Gly Phe Leu Ala Ala Thr Ser Val
20

(2) INFORMATION FOR SEQ ID NO: 515:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq VAVGLTIAAAGFA/GR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 515:

Met Ala Ser Thr Val Val Ala Val Gly Leu Thr Ile Ala Ala Ala Gly
-15 -10 -5

Phe Ala Gly Arg Tyr Val Leu Gln Ala Met Lys His Met Glu Xaa Gln
1 5 10

Val Lys Gln Val Phe Gln Ser Leu Pro Lys Ser Ala Phe Ser Gly Gly
15 20 25 30

Tyr Tyr Arg Gly Xaa Phe Glu Pro Xaa Met Xaa Lys Arg Glu Ala Ala
35 40 45

Gly

(2) INFORMATION FOR SEQ ID NO: 516:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 110 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -83..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq AFSFSRLLSQCRP/DC

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 516:

Met Val Ile Arg Val Tyr Ile Ala Ser Ser Ser Gly Ser Thr Ala Ile
-80 -75 -70

Lys Lys Lys Gln Gln Asp Val Leu Gly Phe Leu Glu Ala Asn Lys Ile
-65 -60 -55

Gly Phe Glu Glu Lys Asp Ile Ala Ala Asn Glu Glu Asn Arg Lys Trp
-50 -45 -40

Met Arg Glu Asn Val Pro Glu Asn Ser Arg Pro Ala Val Gln Gly Pro
-35 -30 -25 -20

His Ala Phe Arg Tyr Lys Ala Phe Ser Phe Ser Arg Leu Leu Ser Gln
-15 -10 -5

Cys Arg Pro Asp Cys Leu Asn Met Leu Arg Arg Phe Ser Gln Tyr Cys
1 5 10

Leu Tyr Leu Val Met Glu Lys Ala Leu Leu Phe Phe Phe Phe
15 20 25

(2) INFORMATION FOR SEQ ID NO: 517:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -42..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 4
seq ITSSLFLGRGSVA/SN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 517:

Met Ser Ser Arg Gly His Ser Thr Leu Pro Arg Thr Leu Met Ala Pro
-40 -35 -30

Arg Met Ile Ser Glu Gly Asp Ile Gly Gly Ile Ala Gln Ile Thr Ser
-25 -20 -15

Ser Leu Phe Leu Gly Arg Gly Ser Val Ala Ser Asn Arg His Leu Leu
-10 -5 1 5

Gln Ala Arg Gly Ile
10

(2) INFORMATION FOR SEQ ID NO: 518:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 55 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Dystrophic muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1

(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq PALCLFDVDGTLT/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 518:

```
Met Ala Ala Pro Gly Pro Ala Leu Cys Leu Phe Asp Val Asp Gly Thr
            -15                 -10                 -5

Leu Thr Ala Pro Arg Gln Lys Ile Thr Lys Glu Met Asp Asp Phe Leu
        1               5                   10

Gln Lys Leu Arg Gln Lys Ile Lys Ile Gly Val Val Gly Gly Ser Asp
15                  20                  25                  30

Phe Glu Lys Val Gln Glu Arg
                35
```

(2) INFORMATION FOR SEQ ID NO: 519:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Normal prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -19..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq ILFHGVFYAGGFA/IV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 519:

```
Met Pro Leu Gly Ala Arg Ile Leu Phe His Gly Val Phe Tyr Ala Gly
                -15                 -10                 -5

Gly Phe Ala Ile Val Tyr Tyr Leu Ile Gln Lys Phe His Ser Arg Thr
            1               5                   10

Leu
```

(2) INFORMATION FOR SEQ ID NO: 520:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq MLLSIGMLMLSAT/QV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 520:

```
Met Leu Leu Ser Ile Gly Met Leu Met Leu Ser Ala Thr Gln Val Tyr
            -10                  -5                   1

Thr Ile Leu Thr Val Gln Leu Phe Ala Phe Leu Asn Leu Leu Pro Val
     5                  10                  15

Glu Xaa Asp Ile Leu Ala Tyr Asn Phe Glu Asn Ala Ser Gln Thr Phe
 20                 25                  30                  35

Asp Asp Leu Pro Ala Arg Phe Gly Tyr
                40
```

(2) INFORMATION FOR SEQ ID NO: 521:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -25..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq WIAAVTIAAGTAA/IG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 521:

```
Met Ser Leu Thr Ser Ser Ser Ser Val Arg Val Glu Trp Ile Ala Ala
-25                 -20                 -15                 -10

Val Thr Ile Ala Ala Gly Thr Ala Ala Ile Gly Tyr Leu Ala Tyr Lys
                 -5                   1                   5

Arg Phe Tyr Val Lys Asp His Arg Asn Lys Ala Met Ile Asn Leu His
         10                  15                  20
```

Ile Gln Lys Asp Asn Pro Lys Ile Val His Ala Phe Asp Met Glu Asp
    25   30   35

Xaa Xaa Asp Lys Ala Val Tyr Cys Arg Cys Trp Arg Ser Lys Lys Phe
40   45   50   55

Pro Phe Cys Asp Gly Ala His Thr Xaa Xaa Asn Glu Glu Thr Gly Leu
                60   65   70

(2) INFORMATION FOR SEQ ID NO: 522:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 108 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -61..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq YTAVSVLAGPRWA/DP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 522:

Met Ser Gly Ser Asn Gly Ser Lys Glu Asn Ser His Asn Lys Ala Arg
    -60   -55   -50

Thr Ser Pro Tyr Pro Gly Ser Lys Val Glu Arg Ser Gln Val Pro Asn
-45   -40   -35   -30

Glu Lys Val Gly Trp Leu Val Glu Trp Gln Asp Tyr Lys Pro Val Glu
                -25   -20   -15

Tyr Thr Ala Val Ser Val Leu Ala Gly Pro Arg Trp Ala Asp Pro Gln
            -10   -5   1

Ile Ser Xaa Ser Xaa Phe Ser Pro Lys Phe Asn Glu Lys Asp Gly His
    5   10   15

Val Glu Arg Xaa Ser Lys Asn Gly Leu Tyr Glu Ile Xaa Asn Gly Arg
20   25   30   35

Pro Arg Asn Pro Ala Asp Gly Leu Asp Trp Trp Ala
                40   45

(2) INFORMATION FOR SEQ ID NO: 523:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 42 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -30..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.9
seq LWMRWTVTSTTRA/WI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 523:

```
Met Ala Ile Ser Leu Arg Ser Ser Gly Ile Ser Val Lys Cys Leu Ser
-30                 -25                 -20                 -15

Lys Leu Trp Met Arg Trp Thr Val Thr Ser Thr Thr Arg Ala Trp Ile
                -10                 -5                  1

Xaa Ala Glu Pro Pro Gln Leu Asp Ile Ser
        5                   10
```

(2) INFORMATION FOR SEQ ID NO: 524:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 65 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -27..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq FVLGSARLGGSGS/MR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 524:

```
Met Ser Glu Val Arg Leu Pro Pro Leu Arg Ala Leu Asp Asp Phe Val
    -25                     -20                 -15

Leu Gly Ser Ala Arg Leu Gly Gly Ser Gly Ser Met Arg Pro Ala Ala
    -10                 -5                  1               5

Met Val Xaa Pro Arg His Gln Gln Pro Pro Leu Leu Pro Asn Gln Leu
```

10 15 20

Pro Ser Leu Leu Arg His Arg Pro Arg Ser Arg Arg Val Arg Thr Ala
25 30 35

Thr

(2) INFORMATION FOR SEQ ID NO: 525:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -15..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LVSATAWLEECWW/SE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 525:

Met Lys Leu Val Ser Ala Thr Ala Trp Leu Glu Glu Cys Trp Trp Ser
-15 -10 -5 1

Glu Leu Ser

(2) INFORMATION FOR SEQ ID NO: 526:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -34..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LYVPLLAVCCLHS/VV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 526:

```
Met Lys Ala Ile Ser Val Ser Leu Leu Arg Leu Thr Lys Leu Leu Trp
                -30                 -25                 -20

Phe Phe Ser Ile Val Leu Tyr Val Pro Leu Leu Ala Val Cys Cys Leu
            -15                 -10                 -5

His Ser Val Val Phe Phe
        1
```

(2) INFORMATION FOR SEQ ID NO: 527:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 143 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Thyroid

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -118..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.8
seq LMIALTVVGCIFM/VI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 527:

```
Met Gly Ser Leu Ser Gly Leu Arg Leu Ala Ala Gly Ser Cys Phe Arg
            -115                -110                -105

Leu Cys Glu Arg Asp Val Ser Xaa Ser Leu Arg Leu Thr Arg Ser Ser
        -100                -95                 -90

Asp Leu Lys Arg Ile Asn Gly Phe Cys Thr Lys Pro Gln Glu Ser Pro
    -85                 -80                 -75

Gly Ala Pro Ser Arg Thr Tyr Asn Arg Val Pro Leu His Lys Pro Thr
-70                 -65                 -60                 -55

Asp Trp Gln Lys Lys Ile Leu Ile Trp Ser Gly Arg Phe Lys Lys Glu
                -50                 -45                 -40

Xaa Xaa Ile Pro Glu Thr Val Ser Leu Glu Met Leu Xaa Xaa Ala Lys
            -35                 -30                 -25

Asn Lys Met Arg Val Lys Ile Ser Tyr Leu Met Ile Ala Leu Thr Val
        -20                 -15                 -10

Val Gly Cys Ile Phe Met Val Ile Glu Gly Lys Lys Ala Ala Gln Arg
    -5                  1               5                   10

His Glu Thr Leu Thr Ser Leu Xaa Leu Glu Lys Lys Ala Arg Leu
                15                  20                  25
```

(2) INFORMATION FOR SEQ ID NO: 528:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -100..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq LASSFLFTMGGLG/FI

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 528:

```
Met Glu Thr Leu Tyr Arg Val Pro Phe Leu Val Leu Glu Cys Pro Asn
-100                -95                 -90                 -85

Leu Lys Leu Lys Lys Pro Pro Trp Leu His Met Pro Ser Ala Met Thr
                -80                 -75                 -70

Val Tyr Ala Leu Val Val Val Ser Tyr Phe Leu Ile Thr Gly Gly Ile
            -65                 -60                 -55

Ile Tyr Asp Val Ile Val Glu Pro Pro Ser Val Gly Ser Met Thr Asp
        -50                 -45                 -40

Glu His Gly His Gln Arg Pro Val Ala Phe Leu Ala Tyr Arg Val Asn
    -35                 -30                 -25

Gly Gln Tyr Ile Met Glu Gly Leu Ala Ser Ser Phe Leu Phe Thr Met
-20                 -15                 -10                  -5

Gly Gly Leu Gly Phe Ile Ile Leu Asp Gly Ser Xaa Ala Pro Asn Ile
                1               5                   10

Pro Lys Leu Asn Arg Phe
        15
```

(2) INFORMATION FOR SEQ ID NO: 529:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 57 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -13..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq MLVLRSGLTKALA/SR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 529:

```
Met Leu Val Leu Arg Ser Gly Leu Thr Lys Ala Leu Ala Ser Arg Thr
            -10                 -5                   1

Leu Ala Xaa Gln Xaa Xaa Phe Ala His Arg Ala Glu Val Arg Lys Ala
     5                  10                  15

Leu Ala Asn Cys Lys Glu Trp Gln Glu Gln Ser Ile Ile Pro Asn Leu
 20                  25                  30                  35

Ala Arg Ile Asp Lys Gln Glu Thr Arg
                 40
```

(2) INFORMATION FOR SEQ ID NO: 530:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Thyroid

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -36..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq NIESLAWTGGTLG/HP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 530:

```
Met Ala Ala Pro Leu Ser Val Glu Val Glu Phe Gly Gly Gly Ala Xaa
    -35                 -30                 -25

Ser Cys Leu Thr Val Leu Arg Asn Ile Glu Ser Leu Ala Trp Thr Gly
-20                 -15                 -10                  -5

Gly Thr Leu Gly His Pro Glu Pro Ala His Leu Asp Gln Glu Glu Phe
                 1                   5                  10

Ala Lys Arg Ala Ala Xaa Val Val His Pro Gly Arg Gln Arg Ala Ala
```

```
        15                  20                  25

Arg Asn Ser Gly Ala Asp Tyr Arg
    30                  35
```

(2) INFORMATION FOR SEQ ID NO: 531:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -65..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq FVGGLPVIFWSWA/GL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 531:

```
Met Thr His Leu Ile Glu Tyr Asp Arg His Arg Lys Ser Arg Leu Ser
-65                 -60                 -55                 -50

Pro Leu Gln His Leu Tyr Leu Leu Pro Ala Asp His Ser Arg Asn Ala
                -45                 -40                 -35

Ala Glu Arg Phe Pro Gly Ala Trp Phe Gln Pro Pro Thr Val Asp Ser
            -30                 -25                 -20

Glu Ala Ser Ala Phe Val Gly Gly Leu Pro Val Ile Phe Trp Ser Trp
        -15                 -10                 -5

Ala Gly Leu Val
    1
```

(2) INFORMATION FOR SEQ ID NO: 532:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 104 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq WARKLLSVPWLLC/GP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 532:

```
Met Ala Ala Ala Ala Leu Gly Gln Ile Trp Ala Arg Lys Leu Leu Ser
        -20                 -15                 -10

Val Pro Trp Leu Leu Cys Gly Pro Arg Arg Tyr Ala Ser Ser Ser Phe
    -5                   1               5                  10

Lys Ala Ala Asp Leu Gln Leu Glu Met Thr Gln Lys Pro His Lys Lys
                15                  20                  25

Pro Gly Pro Gly Glu Pro Leu Val Phe Gly Lys Thr Phe Thr Asp His
            30                  35                  40

Met Leu Met Val Glu Trp Asn Asp Lys Gly Trp Gly Gln Pro Arg Ile
        45                  50                  55

Gln Pro Phe Gln Asn Leu Thr Leu His Pro Ala Ser Ser Ser Leu His
    60                  65                  70

Tyr Ser Leu Gln Leu Phe Glu Gly
75                  80
```

(2) INFORMATION FOR SEQ ID NO: 533:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Hypertrophic prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -38..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq CPLLLLVFTTNNG/RH

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 533:

```
Met Ala Val Glu Ser Arg Val Thr Gln Glu Glu Ile Lys Lys Glu Pro
            -35                 -30                 -25

Glu Lys Pro Ile Asp Arg Glu Lys Thr Cys Pro Leu Leu Leu Leu Val
        -20                 -15                 -10
```

```
Phe Thr Thr Asn Asn Gly Arg His His Arg Met Asp Glu Phe Ser Arg
     -5                  1               5                  10

Gly Asn Val Pro Ser Ser Glu Leu Gln Ile Tyr Thr Trp Met Asp Ala
                 15                  20                  25

Thr Leu Lys Glu Leu Thr Ser Leu Val Lys Glu Val Tyr Pro Glu Ala
             30                  35                  40

Arg Xaa Lys Gly Thr His Phe Asn Phe Ala Xaa Val Phe Thr Asp Val
         45                  50                  55

Lys Arg Pro
     60
```

(2) INFORMATION FOR SEQ ID NO: 534:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -22..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq AVLDCAFYDPTHA/WS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 534:

```
Met Arg Leu Lys Tyr Gln His Thr Gly Ala Val Leu Asp Cys Ala Phe
        -20                 -15                 -10

Tyr Asp Pro Thr His Ala Trp Ser Gly Gly Leu Asp His Gln Leu Lys
     -5                  1               5                  10

Met His Asp Leu Asn Thr Asp Gln Glu Asn Leu Val Gly Thr Met Met
                 15                  20                  25

Pro Leu Ser Asp Val Leu
             30
```

(2) INFORMATION FOR SEQ ID NO: 535:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -86..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.7
seq WAVVLADTAVTSG/RG

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 535:

```
Met Ala Leu Leu Phe Ala Arg Ser Leu Arg Leu Cys Arg Trp Gly Ala
    -85                 -80                 -75

Lys Arg Leu Gly Val Ala Ser Thr Glu Ala Gln Arg Gly Val Ser Phe
-70                 -65                 -60                 -55

Lys Leu Xaa Glu Lys Thr Ala His Ser Ser Leu Ala Leu Phe Arg Asp
                -50                 -45                 -40

Asp Thr Gly Val Lys Tyr Gly Leu Val Gly Leu Glu Pro Thr Lys Val
            -35                 -30                 -25

Ala Leu Asn Val Glu Arg Phe Arg Glu Trp Ala Val Val Leu Ala Asp
        -20                 -15                 -10

Thr Ala Val Thr Ser Gly Arg Gly
     -5                  1
```

(2) INFORMATION FOR SEQ ID NO: 536:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Umbilical cord

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -68..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.6
seq ILLGNYCVAVADA/KK

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 536:

```
Met Ala Ala Ala Ala Ala Gly Thr Xaa Thr Ser Gln Arg Phe Phe Gln
```

-65 -60 -55

Ser Phe Ser Asp Ala Leu Ile Asp Glu Asp Pro Gln Ala Ala Leu Glu
-50 -45 -40

Glu Leu Thr Lys Ala Leu Glu Gln Lys Pro Asp Asp Ala Gln Tyr Tyr
-35 -30 -25

Cys Gln Arg Ala Tyr Cys His Ile Leu Leu Gly Asn Tyr Cys Val Ala
-20 -15 -10 -5

Val Ala Asp Ala Lys Lys Ser Leu Glu Leu Asn Pro Asn Asn Ser Thr
1 5 10

Ala Met Leu Arg Lys Gly Ile Cys Glu Tyr His Glu Lys Asn Tyr Ala
15 20 25

Ala Ala Leu Glu Thr Phe Tyr Arg Arg Thr Gly
30 35

(2) INFORMATION FOR SEQ ID NO: 537:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: kidney

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -60..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq WFYIGSSLNGTRG/KR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 537:

Met Ala Gln Leu Lys Tyr Met Glu Asn Val Gly Tyr Ala Gln Glu Asp
-60 -55 -50 -45

Arg Glu Arg Met His Arg Asn Ile Val Ser Leu Ala Gln Asn Leu Leu
-40 -35 -30

Asn Phe Met Ile Gly Ser Ile Leu Asp Leu Trp Gln Cys Phe Leu Trp
-25 -20 -15

Phe Tyr Ile Gly Ser Ser Leu Asn Gly Thr Arg Gly Lys Arg Val Pro
-10 -5 1

Ala His Phe Ser Asn Thr Ser Leu His Tyr Leu Asn Ala Ala Trp Pro
5 10 15 20

Arg

(2) INFORMATION FOR SEQ ID NO: 538:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -31..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq WSPLSTRSGGTHA/CS

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 538:

```
Met Ser Pro Ala Phe Arg Ala Met Asp Val Glu Pro Arg Ala Lys Gly
    -30                 -25                 -20

Ser Phe Trp Ser Pro Leu Ser Thr Arg Ser Gly Gly Thr His Ala Cys
-15                 -10                 -5                   1

Ser Ala Ser Met Arg Gln Pro Trp
            5
```

---

(2) INFORMATION FOR SEQ ID NO: 539:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 110 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Substantia nigra

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -54..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq SILAQVLDQSARA/RL

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 539:

```
Met Ala Asp Glu Glu Leu Glu Ala Leu Arg Arg Gln Arg Leu Ala Glu
                -50                 -45                 -40

Leu Gln Ala Lys His Gly Asp Pro Gly Asp Ala Ala Gln Gln Glu Ala
            -35                 -30                 -25

Lys His Arg Glu Ala Glu Met Arg Asn Ser Ile Leu Ala Gln Val Leu
        -20                 -15                 -10

Asp Gln Ser Ala Arg Ala Arg Leu Ser Asn Leu Ala Leu Val Lys Pro
    -5                  1               5                   10

Glu Lys Thr Lys Ala Val Glu Asn Tyr Leu Ile Gln Met Ala Arg Tyr
                15                  20                  25

Gly Gln Leu Ser Glu Lys Val Ser Glu Gln Gly Leu Ile Glu Ile Leu
            30                  35                  40

Lys Lys Val Ser Gln Gln Thr Glu Lys Xaa Thr Thr Val Arg
        45                  50                  55
```

(2) INFORMATION FOR SEQ ID NO: 540:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 86 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Cancerous prostate

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -63..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq GLVCAGLADMARP/AE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 540:

```
Met Ser Ala Ala Gly Ala Arg Gly Leu Arg Ala Thr Tyr His Arg Leu
            -60                 -55                 -50

Leu Asp Lys Val Glu Leu Met Leu Pro Glu Lys Leu Arg Pro Leu Tyr
        -45                 -40                 -35

Asn His Pro Ala Gly Pro Arg Thr Val Phe Phe Trp Ala Pro Ile Met
    -30                 -25                 -20

Lys Trp Gly Leu Val Cys Ala Gly Leu Ala Asp Met Ala Arg Pro Ala
-15                 -10                 -5                  1

Glu Lys Leu Ser Thr Ala Gln Ser Xaa Val Leu Met Ala Thr Gly Phe
            5                   10                  15
```

Ile Trp Ser Arg Tyr Ser
20

(2) INFORMATION FOR SEQ ID NO: 541:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 127 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Muscle

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -86..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq TGXLNMTLQRASA/AP

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 541:

Met Ser Asn Tyr Ser Val Ser Leu Val Gly Pro Ala Pro Trp Gly Phe
-85 -80 -75

Arg Leu Gln Gly Gly Lys Asp Phe Asn Met Pro Leu Thr Ile Ser Ser
-70 -65 -60 -55

Leu Lys Asp Gly Gly Lys Ala Ala Gln Ala Asn Val Arg Ile Gly Asp
-50 -45 -40

Val Val Leu Ser Ile Asp Gly Ile Asn Ala Gln Gly Met Thr His Leu
-35 -30 -25

Glu Ala Gln Asn Lys Ile Lys Gly Cys Thr Gly Xaa Leu Asn Met Thr
-20 -15 -10

Leu Gln Arg Ala Ser Ala Ala Pro Lys Pro Glu Pro Val Pro Val Gln
-5 1 5 10

Lys Pro Thr Val Thr Ser Val Cys Ser Glu Thr Ser Gln Glu Leu Ala
15 20 25

Glu Gly Gln Arg Arg Gly Ser Gln Gly Asp Ser Lys Gln Gln Asn
30 35 40

(2) INFORMATION FOR SEQ ID NO: 542:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Uterus

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -18..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq LLGLELSEAEAIG/AD

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 542:

```
Met Ala Asn Pro Lys Leu Leu Gly Leu Glu Leu Ser Glu Ala Glu Ala
            -15                 -10                 -5

Ile Gly Ala Asp Ser Ala Arg Phe Glu Glu Leu Leu Leu Gln Ala Ser
        1               5                   10

Lys Glu Leu Gln Gln
 15
```

(2) INFORMATION FOR SEQ ID NO: 543:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(D) DEVELOPMENTAL STAGE: Fetal
(F) TISSUE TYPE: brain

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -40..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 3.5
seq ALLCTLLLHFQNI/RR

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 543:

```
Met Ile Ile Pro Leu Leu Glu Ile Leu Ile Ile Ile Val Leu Asn Glu
-40                 -35                 -30                 -25

Val Leu Leu Phe Asp Val Asn Ser Val Tyr Lys Ala Leu Leu Cys Thr
                -20                 -15                 -10

Leu Leu Leu His Phe Gln Asn Ile Arg Arg Phe Leu Ser Ser Gln Ser
            -5                  1               5

Pro Met Lys Ala Val Ser Leu Leu Xaa Phe His Gln Pro Asp Phe Asp
```

```
    10                  15                  20

Tyr Ile
 25
```

(2) INFORMATION FOR SEQ ID NO: 544:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 102 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Lymph ganglia

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -52..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7.4
seq LVFIIGLVGNLLA/LV

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 544:

```
Met Asp Ile Gln Met Ala Asn Asn Phe Thr Pro Pro Ser Ala Thr Pro
        -50                 -45                 -40

Gln Gly Asn Asp Cys Asp Leu Tyr Ala His His Ser Thr Ala Arg Ile
    -35                 -30                 -25

Val Met Pro Leu His Tyr Ser Leu Val Phe Ile Ile Gly Leu Val Gly
-20                 -15                 -10                  -5

Asn Leu Leu Ala Leu Val Val Ile Val Gln Asn Arg Lys Lys Ile Asn
                  1               5                  10

Ser Thr Thr Leu Tyr Ser Thr Asn Leu Val Ile Ser Asp Ile Leu Phe
         15                  20                  25

Thr Thr Ala Leu Pro Thr Arg Ile Ala Thr Met Xaa Trp Ala Leu Thr
     30                  35                  40

Gly Glu Ser Glu Met Trp
 45                  50
```

(2) INFORMATION FOR SEQ ID NO: 545:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: AMINO ACID
(D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Homo Sapiens
(F) TISSUE TYPE: Pancreas

(ix) FEATURE:
(A) NAME/KEY: sig_peptide
(B) LOCATION: -29..-1
(C) IDENTIFICATION METHOD: Von Heijne matrix
(D) OTHER INFORMATION: score 7
seq SMIGIGSLPSCWA/CW

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 545:

```
Met Leu Thr Ile Val Lys Ser Pro Gln Lys Ser Tyr Leu Phe Pro Ser
                -25                 -20                 -15

Ser Met Ile Gly Ile Gly Ser Leu Pro Ser Cys Trp Ala Cys Trp Ile
            -10                 -5                   1

Gln Gln Arg
        5
```

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47

A3

(11) International Publication Number: WO 99/06548

(43) International Publication Date: 11 February 1999 (11.02.99)

(21) International Application Number: PCT/IB98/01222

(22) International Filing Date: 31 July 1998 (31.07.98)

(30) Priority Data:
08/905,135 1 August 1997 (01.08.97) US

(71) Applicant (*for all designated States except US*): GENSET [FR/FR]; 24, rue Royale, F-75008 Paris (FR).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): DUMAS MILNE EDWARDS, Jean-Baptiste [FR/FR]; 8, rue Grégoire-de-Tours, F-75006 Paris (FR). DUCLERT, Aymeric [FR/FR]; 6 ter, rue Victorine, F-94100 Saint-Maur (FR). LACROIX, Bruno [FR/FR]; 93, route de Vourles, F-69230 Saint-Genis Laval (FR).

(74) Agents: MARTIN, Jean-Jacques et al.; Cabinet Régimbeau, 26, Avenue Kléber, F-75116 Paris (FR).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report: 8 April 1999 (08.04.99)

(54) Title: 5'ESTs FOR NON TISSUE SPECIFIC SECRETED PROTEINS

(57) Abstract

The sequences of 5'ESTs derived from mRNAs encoding secreted proteins are disclosed. The 5'ESTs may be to obtain cDNAs and genomic DNAs corresponding to the 5'ESTs. The 5'ESTs may also be used in diagnostic, forensic, gene therapy, and chromosome mapping procedures. Upstream regulatory sequences may also be obtained using the 5'ESTs. The 5'ESTs may also be used to design expression vectors and secretion vectors.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PC1, .B 98/01222

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E	WO 98 45437 A (GENETICS INSTITUTE INC. (US); JACOBS K. ET AL.) 15 October 1998 Seq.ID:262 is 97% identical to Seq.ID:38 see page 56 - page 57 see page 63, line 9-17	1-11, 15-37
X	TAKAHASHI N. ET AL.: "Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich alpha2-glycoprotein of human serum" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES USA, vol. 82, April 1985, pages 1906-1910, XP002083674 see page 1908; figure 1	1-37

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 November 1998	08.02.99

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Macchia, G

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/IB 98/01222

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	Database EMBL Emest8, Entry HS1227154 Accession number AA429129 25 May 1997 100% identity with Seq.ID:38 nt.16-64 XP002083675 see the whole document ---	3-8, 15-37
A	YOKOYAMA-KOBAYASHI M. ET AL.: "A signal sequence detection system using secreted protease activity as an indicator" GENE, vol. 163, 1995, pages 193-196, XP002053953 see abstract ---	12,13
A	LIN Y. ET AL.: "Inhibition of nuclear translocation of transcription factor NF-kB by a synthetic peptide containing a cell membrane-permeable moti and nuclear localization sequence" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 270, no. 24, 16 June 1995, pages 14255-14258, XP002050723 cited in the application see abstract ---	14
A	WO 96 34981 A (GENSET (FR); NICOLAEVNA MERENKOVA I.; DUMAS MILNE EDWARDS J.-B.G.) 7 November 1996 cited in the application ---	
A	KATO S. ET AL.: "Construction of a human full-length cDNA bank" GENE, vol. 150, 1994, pages 243-250, XP002081364 cited in the application ---	
A	EP 0 625 572 A (KANAGAWA ACAD OF SCIENCE AND TECHNOL FOUNDATION (JP); KATO S; SEKINE S) 23 November 1994 cited in the application ---	
A	CARNINCI P. ET AL.: "High-efficiency full-length cDNA cloning by biotinylated CAP trapper" GENOMICS, vol. 37, no. 3, 1 November 1996, pages 327-336, XP002081729 cited in the application ---	
A	WO 97 07198 A (GENETICS INSTITUTE INC (US); JACOBS K; MCCOY JM; KELLEHER K; CARLIN M) 27 February 1997 --- -/--	

1

International Application No
PCT/IB 98/01222

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	TASHIRO K. ET AL.: "Signal sequence trap: a cloning strategy for secreted proteins and type I membrane proteins" SCIENCE, vol. 261, 30 July 1993, pages 600-603, XP000673204 ---	
A	HEIJNE VON G.: "A new method for predicting signal sequence cleavage sites" NUCLEIC ACIDS RESEARCH, vol. 14, no. 11, 1986, pages 4683-4690, XP002053954 cited in the application -----	

1

Int. ..ational application No.
PCT/IB 98/01222

INTERNATIONAL SEARCH REPORT

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-37 all partially (Invention 1. on continuation-sheet)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

1. Claims: Invention 1: Claims 1-37 all partially

Nucleic acid comprising the sequence as in Seq.ID:38, complementary sequence, fragments, hybridizing sequences. Polypeptide comprising a signal peptide encoded by said nucleotide sequence. Vector encoding a fusion protein comprising said signal peptide. A method of directing the extracellular secretion of a polypeptide by means of said vector. Method of importing a polypeptide into a cell by means of said signal peptide. A method for making a cDNA encoding a secretory protein, partially encoded by said nucleotide sequence, corresponding cDNA. Polypeptide encoded by said nucleotide sequence, comprising a sequence as in Seq.ID:292, method of making said polypeptide. Method of obtaining a promoter located upstream of said nucleotide sequence, promoter thereof.

2. Claims: Inventions 2-254: Claims 1-37 all partially

Idem as subject 1 but limited to each of the DNA sequences as in Seq.ID:39-291, and corresponding polypeptides, where invention 2 is limited to Seq.ID:39 and 293, invention 3 is limited to Seq.ID:40 and 294,.........., invention 254 is limited to Seq.ID:291 and 545).

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

INTERNATIONAL SEARCH REPORT

:mation on patent family members

Interm' 'nal Application No
PCT/IB 98/01222

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9845437 A	15-10-1998	AU 6956798 A	30-10-1998
WO 9634981 A	07-11-1996	FR 2733765 A	08-11-1996
		FR 2733762 A	08-11-1996
		AU 5982996 A	21-11-1996
		CA 2220045 A	07-11-1996
		EP 0824598 A	25-02-1996
EP 0625572 A	23-11-1994	JP 6153953 A	03-06-1994
		WO 9408001 A	14-04-1994
		US 5597713 A	28-01-1997
WO 9707198 A	27-02-1997	US 5707829 A	13-01-1998
		AU 6712396 A	18-02-1997
		AU 6768596 A	12-03-1997
		CA 2227220 A	06-02-1997
		CA 2229208 A	27-02-1997
		EP 0839196 A	06-05-1998
		EP 0851875 A	08-07-1998
		WO 9704097 A	06-02-1997